



universität
wien

Magisterarbeit

Titel der Magisterarbeit

**“A Discussion of Data Enhancement and
Optimization Techniques for a
Fund of Hedge Funds Portfolio”**

Verfasser

Wolfgang Kappel, bakk.rer.soc.oec.

angestrebter akademischer Grad

Magister (Mag.rer.soc.oec.)

Wien, im August 2008

Studienkennzahl lt. Studienblatt:	A 066 915
Studienrichtung lt. Studienblatt:	Magisterstudium Betriebswirtschaft
Betreuer:	Univ.-Prof. Dr. Dr.h.c. Josef Zechner

Acknowledgements

The topic of this thesis was motivated by my colleagues at Benchmark Capital Management. I want to thank them for the helpful suggestions, discussions and checks for plausibility, especially Karin Kisling, Florian Haberfelner and Olivier Chevaux.

Important advice came from Prof. Josef Zechner, Prof. Thomas Dangl, Michael Halling and Arne Westerkamp. With Ronald Hochreiter I had a lot of very interesting and creative discussions. Thanks a lot!

Many thanks go to my family and friends for support, especially to my mother, Simona, Sandra and Gerhard.

Inhaltsverzeichnis

Acknowledgements	2
1 Introduction.....	10
1.1 German Abstract	10
1.2 English Abstract	12
1.3 Definitions and Setting of the Topic.....	14
1.4 Relevance	16
2 Data Enhancement.....	18
2.1 Hedge Fund specific Problems and Biases	19
2.2 The Database.....	22
2.3 Filtering.....	26
2.4 Illiquidity and Serial Correlation.....	27
2.5 Factor Models.....	33
2.5.1 Introduction.....	33
2.5.2 Risk Factors.....	34
2.5.3 Modeling Dependencies between Risk Factors	36
2.5.4 Non-Linear Dependencies	40
2.5.5 Option Structure.....	41
2.5.6 Principal Component- and Factor Analysis.....	48
2.5.7 Model Selection.....	52
2.6 Cluster Analysis and Qualitative Data	53
2.6.1 Introduction.....	53
2.6.2 Missing Data and (Dis-) Similarity for Qualitative Data.....	55
2.6.3 Cluster Analysis with Return Characteristics.....	57
2.6.4 Individual Peer Indices.....	62
3 Optimization.....	64
3.1 Risk Measures.....	64
3.1.1 Standard Deviation and Mean Variance Analysis.....	65
3.1.2 Sources of Robustness.....	66
3.1.3 Value at Risk	70
3.1.4 Higher Moments.....	74
3.1.5 Other Techniques	75
3.1.6 Conclusion	77
3.2 Limit Architecture.....	77
3.2.1 Linear Constraints.....	78
3.2.2 Logical and Cardinality Constraints	79
3.3 Integration of Data Enhancement and Optimization Techniques.....	81

3.4	Portfolio Optimization Experiments.....	85
3.4.1	General Considerations	85
3.4.2	Quantile vs Dispersion Measures	86
3.4.3	“Unconstrained” Optimization	88
3.4.4	Turnover Constraints.....	91
3.4.5	Fill Up the Track Record.....	96
3.4.6	Conclusion	98
4	Summary.....	99
5	Appendix.....	103
5.1	Black Scholes VBA Code.....	103
5.2	MATLAB Code for Index Resampling.....	104
5.3	MATLAB Code for the Dissimilarity Matrix for qualitative Data.....	107
5.4	MATLAB Code for Stepwise Regression	108
5.5	Metrics for the pdist Function.....	110
5.6	MATLAB Code for Mean – Variance Optimization with a CVaR Constraint	111
5.7	MATLAB Code for Portfolio Optimization with a CVaR and a MAD Constraint	113
	References.....	116
	Curriculum Vitae	121

Table of Abbreviations

2D / 3D	two (three) dimensional
AI	Alternative Investments
ATP	Arbitrage Pricing Theory
AUM	Assets under Management
CAPM	Capital Asset Pricing Model
CARA	Constant Absolute Risk Aversion
CF	Cornish-Fisher
CSFB	Credit Suisse First Boston
CTA	Commodity Trading Advisor
DAX	Deutscher Aktien Index
EDHEC	Ecole Des Hautes Etudes Commerciales du Nord
EUR	Euro
FTR	Fill up Track Record
FX	Foreign Exchange
HFI	HedgeFund Intelligence
HFR	Hedge Fund Research, Inc.
JC	Jaccard Coefficient
LOESS	Locally Weighted Scatterplot Smoothing
MAD	Mean Absolute Deviation
MATLAB	Matrix Laboratory
MILP	Mixed Integer Linear Programming
MVP	Minimum Variance Portfolio
PCA	Principal Component Analysis
PC(s)	Principal Component(s)
S&P	Standard and Poor's
SMC	Simple Matching Coefficient
USD	US Dollar
VaR	Value at Risk

Index of Figures and Tables

Figure 1, Minimum and Maximum Monthly Fund Return	23
Figure 2, Dispersion of Fund Returns	23
Figure 3, Maximum Drawdown	24
Figure 4, Expected Fund Return	24
Figure 5, Higher Moments of Fund Returns	25
Figure 6, Autocorrelation with Indices	28
Figure 7, Autocorrelation with Funds.....	29
Figure 8, Regression Statistics	30
Figure 9, Unsmoothed Returns	32
Figure 10, (Morton, Popova, & Popova, 2005), Formula NORTA Method	37
Figure 11, Original Data.....	38
Figure 12, Random Numbers Generated with the NORTA Method.....	38
Figure 13, Original Data and Random Numbers Generated with Copulas.....	39
Figure 14, Distributions of Strategies Including Options.....	42
Figure 15, Option Payoff – Fixed Costs.....	42
Figure 16, Option Payoff – Black-Scholes Historical Volatility	43
Figure 17, Option Payoff – Black-Scholes Implied Volatilities.....	43
Figure 18, Option Payoff – DAXplus Protective Put	44
Figure 19, Dependency Structures Created with Option Payoffs	45
Figure 20, “Copulas” of Dependencies Created with Option Structures	46
Figure 21, “Copulas” with Empirical Marginals vs “Copula” with Normally Distributed Marginals	46
Figure 22, Pareto Plot of Variance Explained of the First 20 Principal Components	49
Figure 23, (http://en.wikipedia.org/wiki/Cluster_Analysis), Hierarchical Agglomerative Clustering... ..	58
Figure 24, Dendrogram Clustering with Return Characteristics.....	59
Figure 25, Clustered Funds – Correlation vs Expected Return	60
Figure 26, Clustered Funds – Line Plots.....	61
Figure 27, Clustered Funds – Correlation vs CVaR	61
Figure 28, (Lhabitant, 2004, S. 279), Formula Taylor Series Expansion of a General Utility Function.. ..	65
Figure 29, Resampling – Hedge Fund Strategie Weights from 1999 to mid 2007	69
Figure 30, VaR Measures – Empirical, Normally Distributed and Cornish Fisher	70
Figure 31, (Uryasev), Conditional Value at Risk – $CVaR^+$ and $CVaR^-$ are the expected losses strictly (weakly) exceeding the VaR. The CVaR is a weighted average of VaR and $CVaR^+$	72
Figure 32, (Uryasev), Risk Measures	72
Figure 33, (Cornuejols & Tütüncü, 2006, S. 276-277), Formulae of the Linearization of the CVaR 1 ...	73
Figure 34, (Cornuejols & Tütüncü, 2006, S. 278), Formulae of the Linearization of the CVaR 2	74
Figure 35, (Keating & Shadwick, 2002), The Omega Risk Measure.....	75
Figure 36, (Slightly adapted from an Excel Sheet used by Benchmark Capital Management), Omega vs Threshold.....	76
Figure 37, (Chang, Meade, Beasley, & Sharaiha, 2000), Cardinality Constrained Portfolio Optimization	81
Figure 38, Flow Diagram.....	84
Figure 39, Table – Tail Drop Index Level.....	87
Figure 40, Distribution under CVaR Constraint and under Volatility Constraint	87
Figure 41, Table – Tail Drop Fund Level	87

Figure 42, Table Weight Lag	89
Figure 43, Table Time Window.....	89
Figure 44, Cumulative Return HFI Funds.....	90
Figure 45, Table Risk Measures	90
Figure 46, Table Target Return.....	91
Figure 47, Weights Fluctuation with Unconstrained Optimization, CVaR 20%.....	92
Figure 48, Weights Fluctuation with Turnover Constraint, Volatility 10%.....	93
Figure 49, Weights Fluctuation with Turnover Constraint, CVaR 20%.....	93
Figure 50, Weights Fluctuation on Fund Level	94
Figure 51, Table Turnover Constraint Index Level.....	95
Figure 52, Cumulative Return HFR Indices Unconstrained and Turnover Constraint.....	95
Figure 53, Table Turnover Constraint Fund Level	95
Figure 54, Cumulative Return HFI Funds, Fill Up the Track Record.....	96
Figure 55, Table HFR Funds Fill Up the Track Record.....	97
Figure 56, Table Benchmark Funds Fill Up the Track Record	97

1 Introduction

1.1 German Abstract

Ziel dieser Arbeit ist es, einen Überblick über die verschiedenen Techniken zur Datenanreicherung und Optimierung im Falle eines Fund of Hedge Funds Portfolios darzustellen, zu diskutieren und anhand von Experimenten zu illustrieren. Besonderes Augenmerk liegt dabei auch auf der Bewertung des Zusammenspiels der verschiedenen Datenanreicherungs- und Optimierungs-techniken. Erste Bausteine für ein integriertes computergestütztes Anwendungstool werden bereitgestellt und dokumentiert. Zudem werden Ideen für weitere Entwicklungen und Forschung vorgestellt. Zwei wesentliche Punkte unterscheiden diese Arbeit von ähnlichen, nämlich dass sie hauptsächlich auf Fund Level arbeitet und dass sie den gesamten Prozess, beginnend mit der Datenaufbereitung, über die Optimierung bis zur sachgerechten Evaluierung der Ergebnisse behandelt.

Im ersten Teil wird das Thema im Kontext der Finanzwirtschaft verortet, der Begriff Hedge Fund definiert und die Relevanz der Aufgabenstellung erörtert. Neben dem schnellen Wachstum der Hedge Fund Industrie ist besonders das zunehmende Interesse von institutionellen Investoren ein wichtiger Grund quantitative, auf wissenschaftlichen Erkenntnissen aufbauende Methoden zur Unterstützung der Entscheidungsfindung bei der Auswahl von Hedge Funds bereitzustellen.

Der zweite Teil beschäftigt sich mit der Frage der Datenaufbereitung. Generell gilt, dass der Output eines Optimierungs Algorithmus nur so gut sein kann, wie die Qualität der Input Daten mit denen er gefüttert wird. Dies trifft insbesondere auch auf den Fall von Hedge Funds zu, da die Datenlage hier als eher schwierig zu bezeichnen ist: Es werden nur monatliche Renditezahlen zur Verfügung gestellt und Informationen über Risiko Exposures sind nur schwer zu erhalten.

Nachdem ein kurzer Literaturüberblick über die Hedge Fund spezifischen Datenprobleme und Verzerrungen gegeben wird werden die verwendeten Datenbanken anhand von einigen deskriptiven Merkmalen beschrieben. Besonderes Augenmerk wird bei der Datenaufbereitung der hohen Autokorrelation in den Hedge Fund Renditen und dem Auffüllen kurzer Performance Zeitreihen gewidmet. Ersteres weil eine hohe Autokorrelation fundamentalen Prinzipien der modernen Finanzwirtschaft widerspricht, zweiteres weil es zu einer besseren Einschätzung des Risikoprofils der betrachteten Hedge Funds führt.

Zum Zwecke der Datenauffüllung werden im Einzelnen Ansätze über Faktormodelle und Clusteranalyse besprochen. Nach einer Übersicht über die in der Literatur vorgeschlagenen Risikofaktoren wird ein zentraler Gesichtspunkt, nämlich ist die Modellierung von nichtlinearen Zusammenhängen z.B. über Optionsstrukturen, genauer beleuchtet. Wichtige eigene Beiträge in diesem Kapitel sind die ökonomische Interpretation und Motivation des favorisierten Optionsstrukturmodells sowie Vorschläge und erste Experimente zur automatischen Modellselektion und zur Einbindung qualitativer Daten via Clusteranalyse.

Der dritte Teil ist der Optimierung gewidmet. Die Hauptherausforderung ergibt sich aus der Tatsache, dass die Renditen von Hedge Fund Investments meist nicht normalverteilt sind. Da die traditionellen Konzepte der Finanzwirtschaft aber genau auf der Annahme von normalverteilten Renditen aufbauen, müssen alternative Konzepte angewandt werden.

Nach einem kurzen Überblick über die klassische Mean-Variance Optimierung und Möglichkeiten robustere Ergebnisse zu bekommen, werden im Wesentlichen zwei Arten vorgestellt wie mit nicht normalverteilten Renditen umgegangen werden kann: parametrische Ansätze, die die höheren Momente (Schiefe und Kurtosis) der Verteilung berücksichtigen und nichtparametrische Ansätze, die mit historischen oder simulierten Szenarien und den sich daraus ergebenden diskreten Verteilungen arbeiten. Die Präferenzen des Investors können dabei über ein Dispersions- oder ein Quantilsmaß oder einer Kombination aus beidem erfasst werden.

Danach werden Überlegungen angestellt wie einfache lineare und komplexere logische Nebenbedingungen eingesetzt und wie die vorgestellten Konzepte integriert werden können, speziell welche Datenaufbereitungstechniken mit welchen Optimierungsverfahren zusammenpassen. Im letzten Kapitel von Teil drei werden aufwendige Optimierungsexperimente durchgeführt und die neu gewonnen Erkenntnisse interpretiert. Die zentralen Erkenntnisse sind, dass die Wahl des Risikomaßes kaum Einfluss auf das letztinstanzliches Bewertungskriterium, die risikoadjustierte Out-Of-Sample Performance, hat und dass das Auffüllen von kurzen Performance Zeitreihen das Risiko Out-Of-Sample signifikant verbessert.

Abschließend werden die Ergebnisse zusammengefasst und ein Ausblick auf zukünftige Forschungsarbeit gegeben.

1.2 English Abstract

The aim of this thesis is to provide an overview and brief discussion, including some experiments, of techniques for data enhancement and optimization techniques for a fund of hedge funds. Special emphasis is placed on the interaction of the different data enhancement and optimization techniques. First building blocks for a computer based asset allocation tool are provided and documented. In addition it provides some ideas about future development and research. The two main points that distinguish this thesis from papers that treat a similar theme are that it operates on individual fund level and that it covers the whole process beginning with questions of data enhancement and parameter estimation up to proper evaluation of the outcomes.

In the first chapter the theme is put in a broader context of finance, the term “hedge fund” gets defined and the relevance of the problem is reasoned. Besides the rapid growth rates in hedge fund industry the fact that more and more institutional investors invest in hedge funds is an important reason to provide decision support methods based on quantitative models and scientific findings.

The second chapter deals with data enhancement. In general the proverb “garbage in – garbage out” holds true for every optimization algorithm, but it is especially true in the case of hedge funds as the data situation is very difficult in this field: only monthly data is provided and there is only little information about risk exposures.

After a short literature overview about hedge fund specific data problems and biases descriptive statistics are provided for the two databases used in this thesis. With the data enhancement special emphasis is put on the high autocorrelation in hedge fund returns and on filling up track records of funds that are alive for a short time. The former because high autocorrelation is contradictory to fundamental principles of modern finance, the latter because it leads to a better understanding of a funds risk profile.

For the purpose of filling up track records, factor model approaches and the use of cluster analysis are proposed. After a short literature overview about the risk factors considered in literature, the modeling of non linear dependencies, for example via option structures, is discussed on a broader basis as this topic is central in this thesis. Important own contributions in this context are the motivation and economic interpretation of the favored option structure model and some first experiments on automatic model selection and on integrating qualitative data via cluster analysis.

The third chapter talks about optimization. The main challenge is the fact that hedge fund returns are not normally distributed. But as traditional concepts are based exactly on the assumption of normally distributed returns alternative concepts have to be used.

After a short overview of classical mean-variance optimization and possibilities to get more robust outcomes, essentially two alternative concepts are introduced: parametrical approaches, that take higher moments (skewness and kurtosis) into account, and non parametrical approaches, that work with historical or simulated scenarios and with the discrete distributions resulting of these scenarios. With the second approach the preferences of an investor can be captured via a dispersion- or a quantile measure, or a combination of both.

Then, different ways how linear and more complex logical constraints can be used are considered, and procedures how to integrate the concepts presented are discussed, especially which data enhancement and which optimization technique may fit together. In the last part of chapter 3 extensive optimization experiments are conducted and the outcome interpreted. The central findings are that the choice of the risk measure has no significant impact on the out of sample performance, which is the ultimate evaluation criterion; Filling up short track records on the other hand significantly improves the out of sample risk.

Finally the findings are summarized and an outlook for future research is given.

1.3 Definitions and Setting of the Topic

Hedge funds are a subunit of **alternative investments** (AI). There are many different definitions for AI of which a lot do not define it directly, but rather inversely by pointing out that AI are non-traditional investments. But the border of what is considered as traditional is moving as many of the newer investment vehicles get accepted on a broader basis. Other definitions are itemizations of relevant investments. Accordingly AI include besides hedge funds for example managed futures, real estate, commodities and derivative contracts¹, direct investments in private equity, venture and mezzanine capital and leveraged buyouts. But as the number of vehicles grows from year to year, such exhaustive itemizations get antiquated quite fast. Alternatively AI could be characterized according to their return characteristics: “AI are investments that provide unique risk and return properties not easily found in traditional stock and bond investments” (Phillips & Surz, 2003).

In this thesis alternative investments are defined as everything else but long positions in money, stocks or bonds.

To define **hedge funds**, a list of typical features of hedge funds is provided.² Clearly not all of the items have to be the case, but typically they are.

- Hedge funds are allowed to go short and/or to use leverage.
- They require a high minimum investment, often this is 100.000 USD.
- Hedge funds are less regulated than mutual funds, they have to meet only very loose disclosure rules. And as they do not want others to discover their secrets, they have only limited incentives to report voluntarily. As this leads to very low transparency, hedge funds are often considered to be a black box, even for investors that have invested in the fund.
- The manager of a hedge fund has much more freedom in deciding where to invest the money compared to a mutual fund manager, who is typically restricted to a very limited set of assets. This freedom gives the hedge fund manager the opportunity to act in an opportunistic way.
- As hedge funds often invest in illiquid assets and managers need stability in the funding if they want to act flexibly, hedge funds are an illiquid investment themselves. There are several features in the redemption conditions for this purpose: long lock up periods (ranging

¹ See (<http://www.investopedia.com/dictionary/default.asp>).

² The list is based on the introduction chapters of the books of (Lhabitant, 2004), (Lo, 2008), (Stefanini, 2006) and (Gregoriou, 2006) as well as on the experience of fund of hedge fund managers with longtime experience.

from half a year to two years, hard and soft ones), early redemption fees, notice periods, gates (if too many investors want to withdraw their money at the same time).

- For the freedom the investor gives to the manager and for the illiquidity of the investment, the investor expects compensation, such as returns than with mutual funds and/or positive returns also in bad market conditions (“absolute return”).
- Because many hedge funds try to exploit inefficiencies their strategies are often not scalable: If too much capital tries to exploit an arbitrage opportunity, it closes this opportunity by exploiting it. That means that some of the best funds are closed to new investors and that new funds often communicate from the beginning on limits for their assets under management.
- Hedge funds charge very high performance fees. Most of the time the investor has to pay 2% of the investment (management fee) and 20% of the performance (performance fee) each year. On top of this funds of hedge funds normally charge one and ten percent, respectively. These high fees get mitigated a little bit by a high watermark. This means performance fees have to be paid only for gains with respect to the last high. With the rising number of funds and the competition by investable indices and replication strategies there is a trend to lower fees.
- The size of hedge funds in terms of assets under management differs from one million to several billions (the 20% top funds control 80% of assets).³
- Trading strategies are proprietary and therefore must be jealously guarded, even within the organization. This means also that with the departure of a key person the fund is no longer the same one and not seldom the fund gets closed down with the departure – often the general partner *is* the fund.
- Most of the time the general partners own the hedge fund themselves or they invest considerably amount of own money into the fund.
- Many hedge funds are domiciled in tax havens like the Caymans. In terms of assets under management only one third of the funds are onshore.⁴
- Hedge funds have a relatively short lifetime, there is a high turnover. At the high peak in 2005 more than 2000 fund were launched and 850 liquidated.⁵

As in this paper only hedge funds are considered, they are repeatedly simply called funds.

³ (Hedge Fund Research Inc, Q2 2008).

⁴ (Hedge Fund Research Inc, Q2 2008).

⁵ (Hedge Fund Research Inc, Q2 2008).

1.4 Relevance

Numbers for the size of the **industry** differ quite substantially because of in-transparency. Estimations over the last five years differ from 6000 to 9000 funds, with a clear upward trend. At the moment there may exist around 10.000 hedge funds that can be found in one of the hedge funds databases and at least 400 that do not report.⁶ The Economist (2008, March 22nd) talks about quintupled assets held by hedge funds since 2000. The assets under management (AUM) are by now surely over 1 trillion USD (Lo, 2008, p. 1). The recent HFR industry report estimates nearly 2 trillion AUM for Q2 2008. This is still small compared to mutual funds which manage around 14 trillion USD, but the growth rates of the AUM of hedge funds are around 20% compared to 9% for mutual funds.

Besides the size of the industry and the impressive growth rates hedge funds get more and more relevant because of the fact that they are no longer only accessible, interesting and designed for high net worth individuals, foundations and family offices. Inflows come more and more from **institutional investors** like insurances and pension funds. Via investable hedge fund indices the asset class gets accessible even for small private investors.

The focus of this work lies on the needs of institutional investors. They want hedge funds especially for their capability to generate absolute, uncorrelated return. But especially middle and smaller institutional investors do not have the capability to deal with the difficulties that go along with investing in hedge funds. Therefore many of them take the investment via a fund of hedge funds. It is estimated that 16% of the hedge fund assets are administrated by funds of hedge funds.

The **difficulties** arise manly from the in-transparency of the market and methodological questions. The first issue can only get solved by long experience in the market. The focus of this thesis lies on the second issue: As hedge fund returns are highly non-normal distributed and the dependences to risk factors are often non-linear because of the use of derivative instruments and dynamic trading strategies, the traditional concepts of portfolio management cannot be applied. In addition to this one has to be careful when applying new concepts designed to resolve the mentioned problems as they are designed for stock markets with their high frequency data, their high liquidity and their relative stable factor representations.

At the end this thesis wants to give ideas for a software tool for a fund of hedge funds **quantitative analysis**. This tool should be able to prepare data for an optimization algorithm and then run this

⁶ According to a fund of hedge funds manager with longtime experience.

algorithm to get recommendation for an asset allocation. Although such tools already exist, they rely mainly on techniques developed for traditional asset classes and are implemented in a fixed way, so no own ideas can be implemented or tested.⁷ Therefore, given the concept developed after intense literature research and talks to practitioners some methods that can be used in such a tool got implemented in MATLAB and were tested with real data.

⁷ See (<http://www.alternativesoft.com/>) and (<http://www.pertrac.com/per0020/web/me.get?web.home>).

2 Data Enhancement

Although it is not the focus of this thesis, one has to deal with data preparation in any case, particularly with regard to the special data situation for hedge funds. The main aim in this section is to provide the optimization algorithms with the appropriate data. In practice this mainly means to find methods to fill up short track records. Clearly this is a crucial point, as garbage in leads to garbage out.

First we will give a literature overview of hedge fund specific biases. Next we describe the databases we use for our analysis and provide some summary statistics on them. Later, we discuss filtering techniques considered in literature and propose two own filters. The following section deals with serial correlation. We test if this phenomenon exists in our databases and examine the question whether serial correlation is a problem in asset allocation or not. We discuss the unsmoothing approach of (Davies, Kat, & Lu, 2005) and propose directions of further research in this area. Finally we propose a peer group based approach of measuring performance persistence.

The next section (2.5) discusses the usage of factor models to fill up short track records. After briefly discussing other ways to use factor models in the case of hedge funds we summarize and discuss the potential risk factors proposed in literature. Following this, we discuss ways to model the dependencies between the risk factor in a flexible way and illustrate the methods with some experiments using own data. The main challenge is to model non-linear dependencies between funds and factors. Our focus lies on the option structure approach: we propose ways to come up with option prices if no historical data is available and propose motivation and technical reasons why to assume constant prices. Based on literature review we then recommend to use certain option profiles. The last but one topic on factor models is principal component analysis. We introduce this technique, do exploration with own data and propose directions on future research. Finally, we consider problems with model selection.

Another method for filling up track records is developed in section 2.6. First, the usage of a peer group approach and the inclusion of qualitative data are motivated. Next, a method to get a proximity matrix out of the qualitative data and future research topics are proposed. Afterwards, two methods of cluster analysis get introduced: Agglomerative hierarchical and k-means clustering. We illustrate both methods with little experiments using return characteristics of the funds for the

proximity matrix. Finally we propose a procedure to come up with an individual peer index, which can be used to fill up the track record of a certain fund.

2.1 Hedge Fund specific Problems and Biases

When dealing with hedge funds in a quantitative way, there occur some problems that one does not have to deal with in traditional assets classes. The following section discusses these problems and provides some thoughts about how to solve them, if possible.

The first group of problems can be considered as technical problems, which to solve is one important aim of this thesis:

- Hedge funds report their **performance** in general just **once a month**, i.e. there are only 12 data points per year one can work with. This means that all the inter month fluctuations cannot be seen in the data, what may lead to an underestimation of risk when compared to annualized ratios taken from daily return series. One nice example which demonstrates this effect is the behaviour of many quant funds in august 2007, when funds recovered large parts of their losses at the end of the month. In general there is no way to get around this problem; one has to work with monthly data.⁸
- Related to this is the **multi period sampling bias**: analysis often is restricted to funds that have a minimum amount of history available. Clearly quantitative analysis has to have a minimum amount of data to work with, but to set this minimum threshold too high belittles the investment universe/opportunities and restricts the analysis to more established funds, which are said to have different characteristics than younger funds. One way to deal with the problem would be to estimate the parameters for the optimization with a missing value algorithm. But this would lead to severe imbalances between funds: For example a fund who's volatility gets estimated from the one (say very quiet) year it is alive is definitely superior to a similar fund which experienced the last but one (say very turbulent) year. We try to mitigate the problem by filling back track records via regression analysis and clustering approaches. This provides the optimization algorithm with the information about how the fund (modeled as risk factor exposures or via its peers) would have performed under past, maybe more stressful market conditions. For details see 2.5 and 2.6.

⁸ Although most times the investor gets weekly estimates of the net asset value for funds she is invested in, there is no way to include non-invested funds in the analysis on a weekly basis.

- Another main issue is the fact that most of the time returns of the hedge funds are **not normally distributed**. This is a major problem, because “modern” portfolio theory holds only in the case of normally distributed returns or quadratic preferences. As the assumption of quadratic preferences will not be applied in this thesis, alternative ways of portfolio optimizations will be implemented. Section 3 discusses some of these newer methods.
- As hedge funds employ derivative instruments and follow dynamic trading strategies the dependencies to the underlying risk factors are often **not linear**. In addition to this the fact that hedge fund managers often act very opportunistic in changing market conditions causes very unstable factor model representations. Sections 2.5.4 and 2.5.5 discuss these problems.
- Hedge Funds often have a short lifetime i.e. they have a very **short track record**. This can mean that the fund did not experience stressful market events, and this may underestimate risk substantially, as no one knows how the fund performs in stormy market conditions. The distributions, correlations and what else estimated or fitted from this short track record will not be accurate. How to fill up the missing history for such a fund will be treated in sections 2.5 and 2.6.
- An effect that is more a “risk bias” is the one that a quite high **serial correlation** can be observe in the returns of some hedge fund (strategies). Section 2.4 discusses ways to handle this problem.

The second group of problems are all in some sense “return biases”⁹ that arise from problems with the database used. But as we consider only hedge funds in our analysis these database biases are not so important for us as we do not have to compare with assets that are not affected by the below mentioned biases. Nonetheless these biases should be kept in mind as the outcome of portfolio optimization experiments is likely to be effected by some of these biases in the sense that the same experiment setting will lead to different conclusions when applied to a biased and a unbiased sample.

- Related to the topic of short track records are the so called the **instant history** bias, as this bias occurs when fund managers try to provide longer track records for their new funds. The

⁹ The most heavily cited papers about biases in the hedge fund world are (Fung & Hsieh, 2002) and (Fung & Hsieh, 2000).

bias occurs if a fund enters a database with a track record built up in an incubation time, were the fund managed a little amount of money from the manager' friends or the company mother. If this track record is good, the fund reports it to the database, if not the track record will not be reported and another of this test funds will be created. This biases accounts for 1.4% of overestimating hedge fund returns, according to (Fung & Hsieh, 2000). To get around this bias one can just delete the first year track record or replace it by own estimates for this period.

But there may be other reasons for the fact that a fund's return is higher in the first year than in the following years: often hedge fund managers come from prop trading desks of big investment banks and after some years of successful trading there they decide to start up an own fund. But at the prop desk they had all the information sources and connections from the investment bank. For some time they are able to feed on these sources but as time goes by they get cut off from the hot info.¹⁰ Or it may be that the managers detected a new arbitrage opportunity they could exploit in the beginning and then this opportunity closes.

- The **survivorship bias**, which occurs if only alive funds are taken into account, is more relevant in the case of hedge fund indices and when comparing hedge funds to other asset classes. Here it's estimated that this bias accounts for up to 3% overestimating hedge fund returns (Fung & Hsieh, 2000). For this paper it is not so relevant, as the optimization horizon is just one month and we only consider hedge funds in the optimization. If it happens that a fund selected by the optimization algorithm at time t stops reporting exactly at $t+1$ one can set its performance to a defined number in the evaluation. The problem with choosing this number now is, that we don't know if the fund stopped reporting because
 - it was liquidated or
 - it decided to stop the (voluntary) reporting because the performance was so bad that it would have negative impact on the marketing for other funds from the same company or
 - it stopped reporting because the performance was so good that it does not have to attract new capital anymore.

Clearly there is a higher chance that the fund got liquidated if it had negative returns and/or capital outflows just before it stopped reporting. For details how to estimate the probability of liquidation see the paper by (Horst & Verbeek, 2007).

¹⁰ According to a fund of hedge funds manager with longtime experience.

- Another fact that can be named a “return bias” is that most of the hedge funds are **denominated in USD**, which makes it necessary for an EUR investor to either take the currency risk into account in the optimization or to hedge the currency risk. In this thesis only use funds that are denominated in USD will be used for the calculations. Clearly here would be room for improvement.

2.2 The Database

In this study two different databases are used. The first is the **in-house database** of Benchmark Capital Management GmbH. It's well maintained and contains over 750 funds and more than 250 indices. For all funds there are monthly returns reported, for some of them detailed qualitative data is available. One nice thing about the Benchmark database is that it is free from survivorship (dead funds stay in the database) and instant history bias (performances are checked before they enter the database).

The second database is much broader. It's a spreadsheet (Excel) image of a commercial hedge fund database provided by Hedge Fund Intelligence (**HFI**) and it contains more than 4500 funds, out of which 711 haven a focus on asia, 1782 in Europe and the rest (2034) in the US. Monthly returns of the funds and a set of qualitative data are available for all funds. For the HFI database we have data from June 1998 to March 2008 available. In the Benchmark database the first entries start in 1990, but to get comparable results and because for the early and mid 90ies there are only a few funds we use the same time period as with the HFI database for our calculations.

Unfortunately the HFI database is not free of survivorship bias therefore comparisons between the two databases have to be interpreted with caution.

To give an idea about the characteristics of the two databases some summary statistics are provided, but instead of printing a table with a lot of numbers we rather plot histograms with the distribution of the characteristics among the funds. The blue bars are for the Benchmark database funds and the red ones for the HFI database.

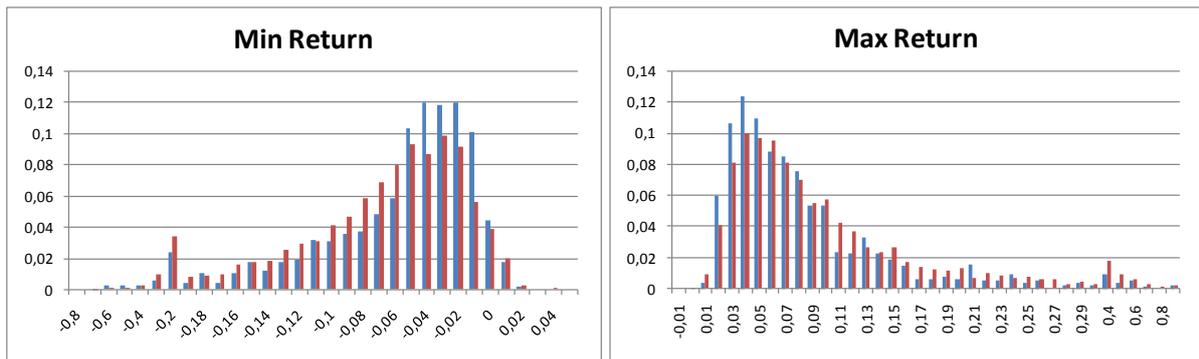


Figure 1, Minimum and Maximum Monthly Fund Return

The first two figures show the distribution of the minimum and the maximum monthly return of the funds. It can be seen that a large fraction of funds face losses of over 10% per month at least once in their lifetime (B: 19%, HFI: 26%), losses of over 20% occur for a considerable fraction of funds (B: 4%, HFI: 5%). The mean minimum return is (B: -6.6%, HFI: -7.7%), the median is (B: -4.7%, HFI: -6.1%). The worst losses were (B: -70%, HFI: -78%). On the other hand also the upside is there: (B: 24%, HFI: 31%) of the funds had a maximum monthly return of over 10% and (B: 8%, HFI: 10%) over 20%. The absolute best monthly returns are impressive (B: 83%, HFI: 254%). Mean and median are (B: 8.4%, HFI: 9.9%) and (B: 6.1%, HFI: 6.9%), respectively. It seems that the Benchmark database contains more defensive funds than the broader HFI database. This impression is confirmed if one looks at the next figure that shows two measures of dispersion.

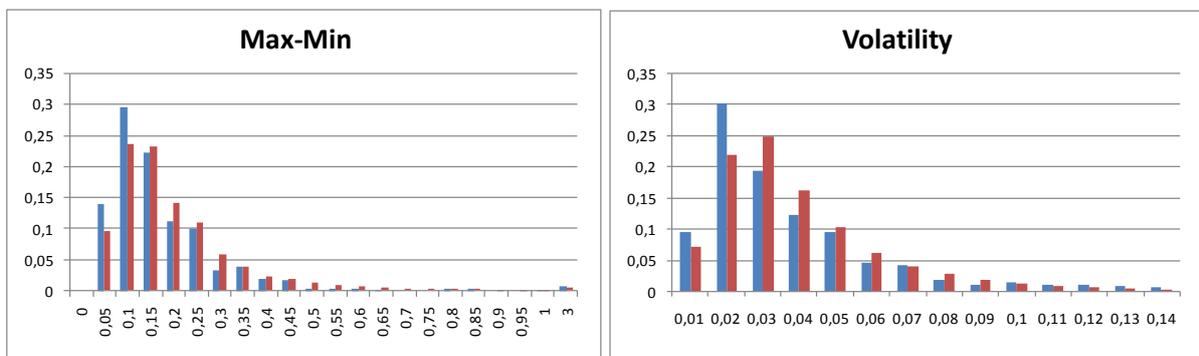


Figure 2, Dispersion of Fund Returns

In both cases the mean and the median are lower for the Benchmark funds:

	Max-Min	Volatility (annualized)
mean	(B: 15.1%, HFI: 17.6%)	(B: 9.4%, HFI: 12.1%)
median	(B: 11.3%, HFI: 13.6%)	(B: 7.3%, HFI: 9.6%)

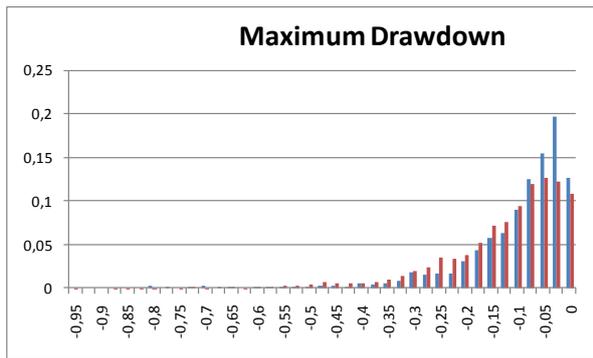


Figure 3, Maximum Drawdown

A similar pattern can be seen in the distribution of the maximum drawdown, the Benchmark funds have a lower mean (B: 11.2%, HFI: 14.2%) and median (B: 7.8%, HFI: 10.6%)

Nonetheless the expected annualized return is higher with the Benchmark funds (B: 12.5%, HFI: 11.7%), median (B: 10.6%, HFI: 9.5%):

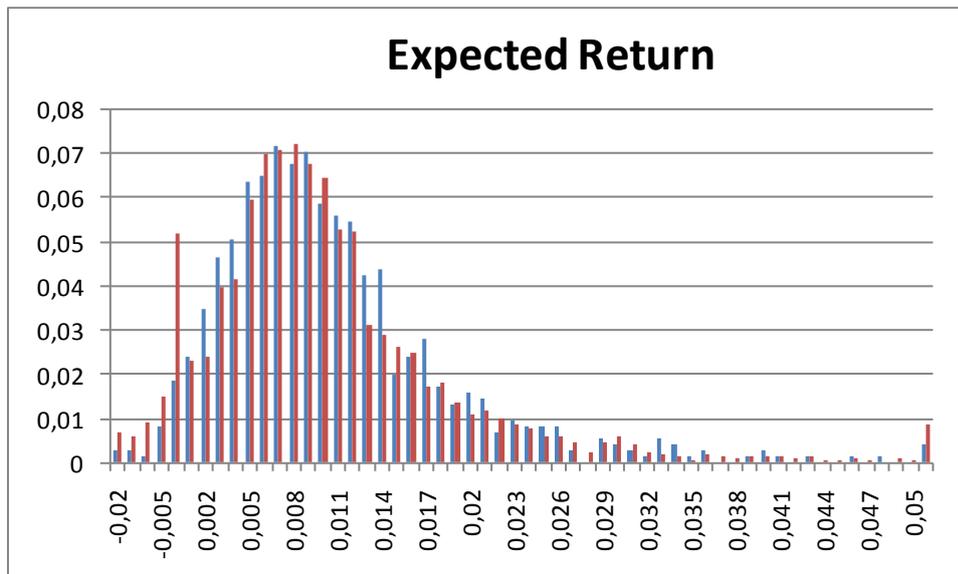


Figure 4, Expected Fund Return

But there may be a bias in the Benchmark database: as mainly funds with an impressive past track record are considered for further analysis and as they are included with their full (live) track record in the Benchmark database it is likely that the database contains a considerable upward bias.

If one looks at the distributions of the skewness and the kurtosis of the funds one could have the impression that the coefficients just fluctuate randomly around zero and that for the hedge fund industry as a whole the assumption of normally distributed returns may be justified. One may argue that the fluctuations mainly come from the fact that only a little number of observations is available to estimate the moments. Indeed the mean and the median are close to zero for the skewness in

both database: (B: -0.17%, HFI: -0.1%) and (B: -0.18%, HFI: -0.11%), respectively. Also with the kurtosis the values are relatively close to zero:¹¹ For the mean we observe (B: 3.3%, HFI: 2.4%) and (B: 1.6%, HFI: 1.1%). The mean is in this case not the best summary as the kurtosis is (similar to the volatility) right skewed by construction and high outliers like 107 (the maximum observed kurtosis in both databases) bias the mean upwards.

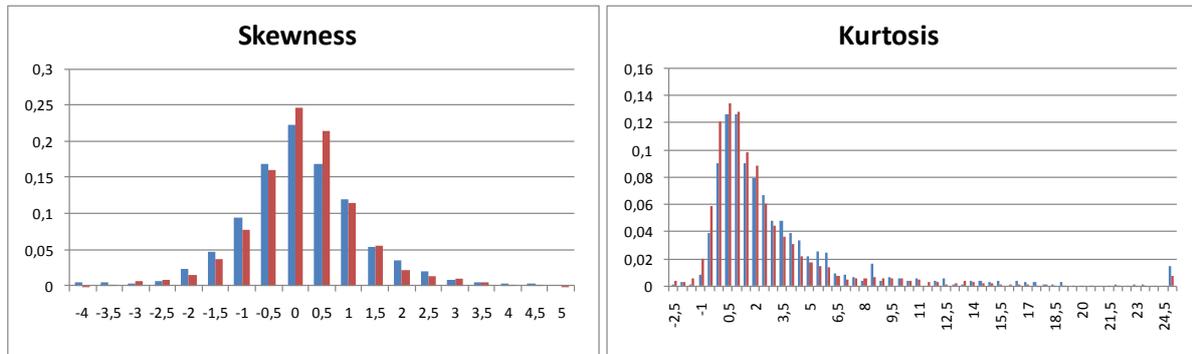


Figure 5, Higher Moments of Fund Returns

But the impression from the figures is wrong. First also hedge fund indices show negative values of skewness and high positive values of kurtosis, even indices that aim to cover the whole industry like the HFR Fund Weighted Composite (-0.6, 3), the HFR Fund of Fund (-0.27, 4.23), the HFI Absolute Return Composite (-0.49, 1.44) and a simple average index of the Benchmark database funds (-0.46, 3.67). From our point of view one explanation for this phenomenon can be that the large negative values in certain common months that account for the negative skewness in the left skewed funds cannot be offset by the return of the funds with positive kurtosis during *these* months. The positive skewness funds may have no severe drops in *these* months (otherwise their own skewness would be negative, too) but they seem to face also small drops, at least they do not offset the severe drops of the left skewed funds by high gains. In other words the left tail correlation among positive and negative skewness funds must be positive, otherwise the central limit theorem would lead on the aggregate level to a normal distribution.

The second argument is that for 58% of the funds in the Benchmark database and for 46% in the HFI database the hypothesis of normally distributed returns is rejected by the Jarque-Bera-Test at a 95% level. Still 44% and 34%, respectively, are rejected even on a 99.9% level. As the Jarque-Bera-Test account for the number of observations in our calculations on fund level and as on index level relatively long time series are available the argument that non normality comes from short track records does not hold true either.

¹¹ We calculated the excess kurtosis.

2.3 Filtering

The first question is to work on **fund** or on **index** level. For this thesis we decided to work mainly on fund level. This distinguishes this thesis from other works that deal with optimization in the case of hedge funds because these papers mainly just use hedge fund indices for their calculations.

One way to perform the optimization would be to use all the funds available in the database. However there are ways and good reasons to filter funds before doing the data enhancement and the optimization. The following filter approaches are considered in literature:

- In most of the databases there are funds denominated in different **currencies** available. If one does not want to deal with currency risk, one has to work only funds that report their returns in one certain currency, normally in USD. We will also do so in our studies.
- Another filter frequently used is to restrict the sample to funds that have a certain minimum amount of **assets under management**.
- One natural way to filter funds is to rank them by their **alpha** and then just use the best say 100 funds for the optimization. It is even possible to get an asset allocation decision by allocating $1/n$ out of the n best funds from the alpha ranking. The challenge here is to determine the alpha. For this question see 2.5. (Alexander & Dimitriu, 2005) focus on the filtering according to alphas and conclude that the alphas from different models, although different in their values, lead to a similar outcome in terms of the ranks. The out of sample performance of their alpha ranked minimum variance portfolios is superior to an equally weighted portfolio of all funds and to that of randomly selected MVPs.
- Another approach is to rank the funds according to their **Hurst exponent**. The Hurst exponent can be considered as a measure of persistence or autocorrelation. (Amenc, Bied, & Martellini, 2003) use it just to demonstrate that there is persistence in hedge fund returns and then focus on predictability in lagged economic variables for hedge fund returns. As a filtering tool it is used only in (Olszewski, 2005). But as the outcomes described there are not very convincing we will not use the Hurst exponent as a filtering tool. How we will address for autocorrelation will be described in section 2.4.
- There is also a way to filter funds driven by more technical considerations. As with the mentioned problems with hedge fund quantitative data it may be a good idea to just work

with funds that have a **sufficient amount of data** (attention, this lead to the so called multi period sampling bias)

In addition we want to propose the following filters:

- A more qualitative filter is to use only funds that are already invested and such that have passed the **due diligence** process and were considered as investable or eligible. This can be a first and easy way to get qualitative information into the optimization process.¹²
- Only use funds for quantitative analysis that show strong and **reliable patterns** in the factor model representation or that are good to assign to a peer group in the cluster analysis – in other words: funds that are relatively easy to handle in a quantitative way. Out of these funds one can construct a core portfolio about which we can make statements with relatively high certainty. Around this portfolio the newer funds and such that follow very exotic strategies can be allocated by hand.

2.4 Illiquidity and Serial Correlation

Most of the problems one has to deal with in the context of hedge funds are more of a technical kind, like going away from the assumptions of normality and linearity. Serial correlation on the other hand is more an **economic problem**: it violates the Random Walk Hypothesis; with the presence of autocorrelation there is unexploited predictability in the data; (spurious) serial correlation leads to underestimation of beta coefficients, overestimates the risk adjusted performance as measured by the Sharpe Ratio for example and it leads to wealth transfers between new, existing and departing investors.

The problem with predictability in returns is that this implies that the manager's investment policy is not optimal: if there is a higher probability for his returns to be positive next month, he should take advantage of this forecast and increase his positions and vice versa.¹³ A counterargument could be that the strategy the manager follows is not scalable or that there are other restrictions that prevent an increase in the position. Normally other market participant then would take advantage of this opportunity, but maybe there are reasons that prevent them from doing so. In this case **market**

¹² One problem we faced with this approach is that we do not have information about when the fund entered the investable universe. As it is very likely that only funds that performed well over a longer period of time enter we may import a sustainable bias.

¹³ For more detailed explanations see (Lo, 2008, p. 64).

inefficiencies like barriers to entry, illiquidity or lack of information would be the cause of the serial correlation and the reason why the market is not able to arbitrage it away.

(Lo, 2008) argues in his recent book that serial correlation is most likely the result of **illiquid securities** that are contained in the fund. This is consistent with our argument as illiquidity forestalls quick changes in positions. But this slow adaption may explain serial correlation only to a smaller extent as the market would find ways to speed up adaption to exploit the profit opportunity if it is too attractive. A more important explanation for serial correlation caused by illiquid instruments is the lack of market prices. Prices then get extrapolated from older prices always using some kind of smoothing profile. In particular (and in addition) (Getmansky, Lo, & Makarov, 2004) talk about “linear extrapolation of prices for thinly traded securities, the use of smoothed broker-dealer quotes, trading restrictions arising from control positions and other regulatory requirements, and, in some cases, deliberate performance-smoothing behavior” and stale pricing resulting from large non-trading intervals. Empirical support for the illiquidity theory is given by the fact that “funds with the highest serial correlation tend to be the more illiquid funds, e.g., emerging market debt, fixed income arbitrage” (Getmansky, Lo, & Makarov, 2004).

The question now is: is it necessary to deal with serial correlation in the investment process and if yes, what to do about it? First we will test if the phenomenon of serial correlation can be found in the databases we use, too. We find the following distributions of serial correlation¹⁴:

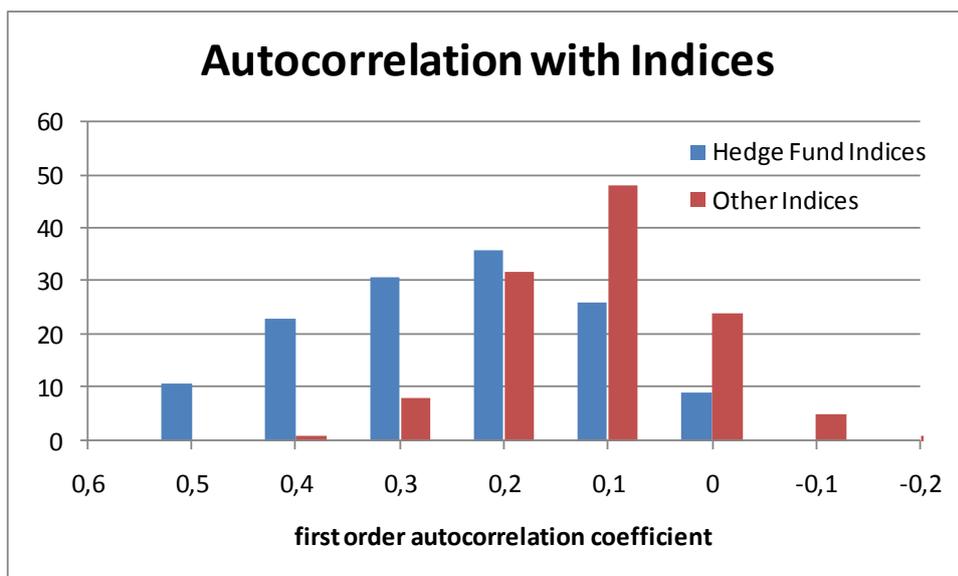


Figure 6, Autocorrelation with Indices

¹⁴ The autocorrelation was calculated over the whole lifetime of the index or the fund respectively. On fund level we only considered funds with a minimum amount of 36 observations.

The figure clearly shows that hedge fund indices show more positive serial correlation than other indices like stock market-, regional-, commodity and currencies indices. Especially the number of indices that show autocorrelation coefficients that are significant different from zero at a 95% level (autocorrelation larger than approximately 0.2)¹⁵ is much larger in the hedge fund universe. The next figure shows the distribution of the autocorrelation coefficients for individual hedge funds.

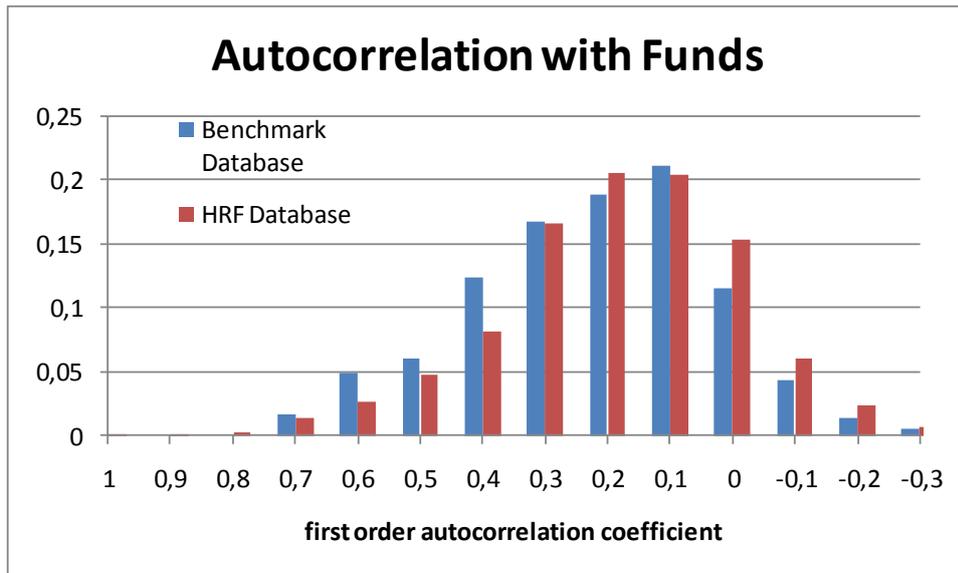


Figure 7, Autocorrelation with Funds

Here we also observe positive autocorrelation. As the funds have very different lifetimes an overall confidence level would be too imprecise. Therefore we computed the percentage of funds that have autocorrelation coefficients significantly different from zero. In the Benchmark database 37% show significant serial correlation, in the HFI database 28%. This supports the optical impression that the Benchmark database is somehow biased to more illiquid funds.

The second question to ask is whether serial correlation is a problem for an optimization algorithm or not. Clearly most optimization techniques will favor funds with high positive autocorrelation as they have lower volatility with same expected return and lower beta coefficients to the overall market as their counterparts with no autocorrelation.

But is this a bad thing? If the picked funds behave similar to the risk free rate but with a higher expected return there would be no problem. But, as with risky bonds for example, one could assume

¹⁵ This number was calculated a little bit inaccurate because of the fact that the indices have different lifetimes and the level of significance is calculated by $1.96/\sqrt{V}$ (average number of observations per index).

that the higher expected return is compensation for a higher default risk. In the case of hedge funds it could also be a higher probability of a sharp drop in the net asset value. This sharp drop can be for example due to a realized loss with an investment object that was in the books for years with nicely increasing linear extrapolated prices.

The assumption that high autocorrelation goes along with high tail risk would also be consistent with the finding that many hedge funds have option like returns, especially similar to short put options: under normal market conditions these funds produce high stable income but if the market gets hit they get hit even more. Factor models with option structure will be discussed in section 2.5.5.

Here we will just perform a short check and test whether there is a relation between serial correlation and the probability of a **sharp drop**, defined as a 3-sigma or worse event. We regressed the number of 3 sigma events on the dummy variable significant autocorrelation. This time we include all funds, also the one with less than 36 entries as it is not unlikely that a fund gets liquidated after such an event even if it is very young.

And indeed we find a significant relationship between the existence of significant autocorrelation and the number of 3 sigma events for all the hedge fund samples: the hedge fund indices, the funds in the Benchmark database and the funds in the HFI database. Only for the traditional indices we could not find a significant relationship, but this is due to the fact that there is hardly any significant autocorrelation. The following table shows the statistics for the pool of all hedge fund indices.

Dependent Variable: THREESIGMA

Method: Least Squares

Date: 09/05/08 Time: 11:01

Sample: 1 138

Included observations: 138

	Coefficient	Std. Error	t-Statistic	Prob.
C	0.363056	0.097164	3.736513	0.0003
AUTOCORREL	1.195339	0.392463	3.045737	0.0028
R-squared	0.063854	Mean dependent var		0.608696
Adjusted R-squared	0.056971	S.D. dependent var		0.655519
S.E. of regression	0.636572	Akaike info criterion		1.948949
Sum squared resid	55.11050	Schwarz criterion		1.991373
Log likelihood	-132.4775	Hannan-Quinn criter.		1.966189
F-statistic	9.276516	Durbin-Watson stat		1.955776
Prob(F-statistic)	0.002788			

Figure 8, Regression Statistics

If there is risk inherent in auto correlated returns, calculating performance statistics like the Sharpe ratio out of the observed data can lead to serious misspecifications. (Getmansky, Lo, & Makarov, 2004) propose ways to calculate adjusted Sharpe ratios. In an optimization framework illiquidity can be taken into account via a limit on the illiquidity, among other. Illiquidity can be measured for this purpose as the autocorrelation of the funds or as the liquidity terms of the fund, like the lock up in months.¹⁶

Another way to deal with autocorrelation is to “**unsmooth**” the return series. (Davies, Kat, & Lu, 2005) propose the following procedure:

$$r_t = \alpha^{-1} (r_t^* - (1 - \alpha)r_{t-1}^*)$$

where r_t is the unsmoothed return and r_t^* is the observed smoothed return and they propose to set $1 - \alpha$ equal to the first order autocorrelation coefficient. The resulting time series r_t has the same expected return as the observed one, but a higher volatility and no autocorrelation. This procedure implicitly assumes that the hedge fund managers use a single exponential smoothing approach of the form

$$r_t^* = \alpha r_t + (1 - \alpha)r_{t-1}^*$$

The economic intuition why fund managers are assumed to exactly use a single exponential smoothing is not of interest in our context, the procedure will be simply used to increase the risk of autocorrelated funds. A problem that could occur with this is if the unsmoothing procedure is only applied to funds with significant autocorrelation, the funds with high, but insignificant autocorrelation would have an unjustified advantage and there could emerge spurious dependencies. Therefore the procedure has to be applied to all funds.

Another problem with the unsmoothing of the return series is that the risk that cannot be seen in the smoothed return series, but that is assumed to be there, gets redivided over the whole observation period, as can be seen below. But as argued before it is more likely that this risk will appear as a sharp drop. To model this would require a simulation approach and further research about when to assume this sharp drop.

¹⁶ It is a good guess that funds with long lock up and notice periods will also have a high illiquidity exposure.

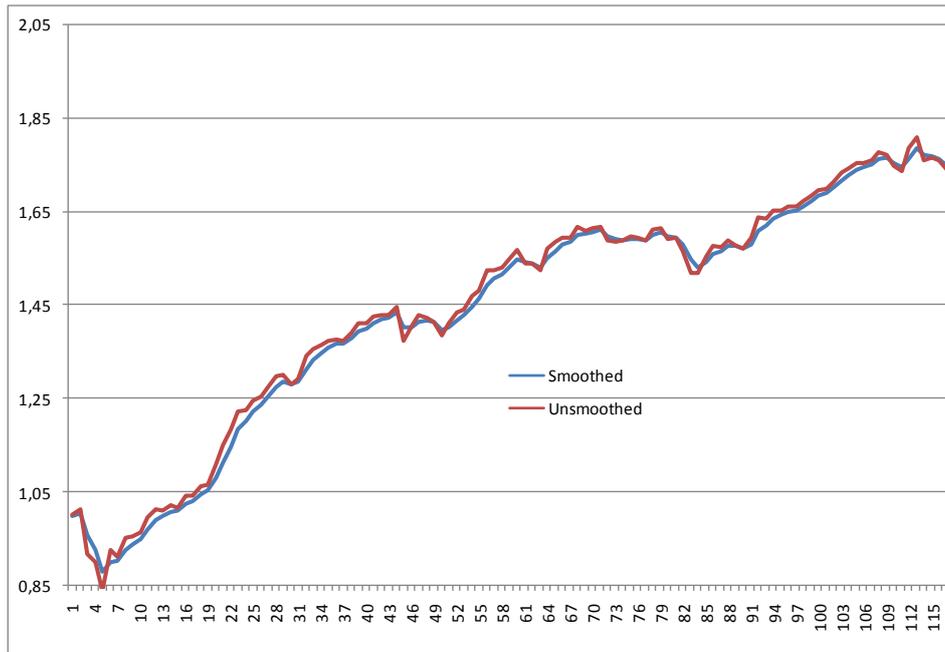


Figure 9, Unsmoothed Returns

Whether this procedure improves the out of sample performance of the optimization algorithm or not will be shown at the end of section 3.4.3. To give a short preview: it does not!

Another comment on serial correlation: It is closely related to performance persistence¹⁷ (which usually implies positively auto correlated returns) and significant beta coefficients for lagged index returns¹⁸, two phenomena that are documented in the academic literature.

In an in-house study (Haberfelner & Kappel, 2007) performed a test for performance persistence with the following design: the measure for performance was “number of months in the **top quartile** of a peer group in terms of the Sharpe ratio during one year”. Then the influence of the fact that a fund was more than 9 months in the top quartile in year $t-2$ ¹⁹ on the probability of again being in the top quartile for more than 9 months in year t was measured: For an average fund the probability of being a top performer in the above sense was a little below 10%, for a fund that was a top performer already in $t-2$ it was above 20%. The nice thing with this study is, that it uses a relative performance measure, comparing funds within a homogenous peer group, all acting on the same markets, using similar strategies and therefore exposed to the same risks. Our measure is therefore better in capturing manager skill than an alpha where important sources of risk may not be taken into

¹⁷ See for example (Agarwal & Naik, 1999).

¹⁸ See for example (Asness, Krail, & Liew, 2001).

¹⁹ The reason why we had to take two years difference was because the rolling Sharpe ratio was calculated with a one year window.

account. With this relative robust measure of manager skill we found performance persistence across all strategies, in the liquid ones as well as in the illiquid ones.

2.5 Factor Models

2.5.1 Introduction

As mentioned in section 2.1, one of the most severe problems when dealing with hedge funds in a quantitative way is the limited amount of data. If the analyst limits her analysis to funds that have a longer **track record** (say over 48 data points) she is likely to import the above described multi period sampling bias. We propose two different ways to deal with this problem: the first one is based on regression, the second one on cluster analysis.

The idea behind the regression approach is to find a factor model representation out of the data we have available for the fund and then calculate the predicted return for the fund out of the known past returns of the factors. We will use some “tricks” to model nonlinear dependencies between fund and factor. The factor model representation gives us an idea how the strategy the fund seems to follow according to our analysis would have performed under stressful past market events.

The second approach takes account of the problem that factor model representations may not be stable over time. Here we do not try to find risk factors that explain the funds return but we try to find funds that are similar to the one we observe but have a longer track record. Out of the n - most similar funds we try to find characteristics for the observed fund and write back its track record. This method will be discussed in section 2.6 in detail.

But factor models can be used not only for filling up the track record of a fund. Another application can be to use not the actual but the risk adjusted return as an input for the optimization algorithm. In addition one can use factor models to determine its alpha and to check its strategy or regional classification. We will discuss these additional applications in brief beneath.

Regression analysis can be used to determine the **alpha** of a fund as a measure of manager skill. (For a more sophisticated approach the analyst can also look for patterns in the residuals – at the end the alpha *and* these patterns are what the investor pays the high fees for). By using hedge fund specific factors it is even possible (to some extent) to distinguish between “true alpha” (real manager skill, exploiting arbitrage opportunities and inefficiencies) and taking “non mainstream risks” that looks

like alpha in a regression with traditional risk factors. One can use the alpha to filter funds before the optimization algorithms get applied and to construct a simple $1/n$ portfolio out of the n “alpha best” funds as one possible method to come up with an asset allocation.

Maybe even more useful regression analysis is in **checking** if the fund manager really does what he says he’s doing; in checking if the fund is classified right; in detecting non-reported style drifts and changes in leverage via rolling regressions. But as this application is used in a more in depth analysis of a single fund, we will not discuss it here in detail.

2.5.2 Risk Factors

The problem with factor models like the one of Sharpe²⁰ and especially when used with traditional indices, is that they were designed to explain the return of long only managers of mutual funds. Clearly linear factor models with long only indices of equity and bond markets as regressors will not be able to explain significant proportions of hedge fund returns, that are generated using OTC-derivatives, bets on firm and macro events, leverage, short positions and a lot of other trading strategies.

On the other hand it is also clear that no model will ever be able to fully explain the well protected secret strategies of hedge fund managers. Nonetheless there exist ways to at least better explain hedge fund returns as with the traditional factors and models. We will now discuss potential risk factors considered in literature to explain hedge fund returns.

A first step on the **equity** side could be to include the factors proposed by (Fama & French, 1992). These factors are HML (high minus low book to market ratio), SMB (small minus big firm size). Another well known factor in academic literature is the one proposed by (Carhart, 1997), the momentum factor or WML (winner minus losers). (Chen, Roll, & Ross, 1986) propose unanticipated changes in industrial production and in inflation, as these factors drive the expected dividend payments. Many hedge funds, especially the ones that act in emerging markets, are specialists for one particular **region** or country. Therefore it will be useful to include regional indices, too. MSCI provides us with a large number of regional indices, even subdivided into value and growth indices.²¹

²²

²⁰ See (Sharpe, 1988) or (Sharpe, 1992).

²¹ See (<http://www.msibarra.com/>).

²² The link between value/growth and the Fama-French/Momentum factors is that an aggressive growth portfolio will mainly contain small size stocks with low book to market ratios and high momentum. Confer

Other factors could be the dividend yield on the S&P500 as a proxy for time variations in the unobservable risk premium, sector indices (like the ones provided by MSCI), traded volume as a proxy for liquidity (NYSE cumulated volume), implied volatility indices (like the VIX for options on the S&P 100)²³ as a measure of the market's expectations about future volatility and the intra month volatility on equity and bond markets as a measure for current stability. As rolling volatility is highly auto correlated by construction it is better to use changes in volatility instead. From our point of view volatility is especially relevant for hedge funds as it is a mayor input for all kinds of hedging instruments like options for example and it therefore determines significantly the cost of many strategies applied by hedge funds.

On the **bond** side useful factors, besides the traditional ones like short, mid and long term government bond and corporate bond indices, could be: The term spread (difference between long term government and treasury bond yields) and the default yield spread ("Baa and under" bond return minus long term government bond). These factors are not only useful because fixed income hedge funds take bets on them, but also because they influence the discount rate in a discounted cash flow framework for stock prices.²⁴ And the yield on the 3-month treasury bill rate is negatively correlated with future stock market returns, thus it can serve as a proxy for the expectations about the economic outlook.²⁵ In addition the cost of capital is essential for many hedge fund strategies as many of them need high leverage to be profitable.

Especially global macro hedge funds and CTAs (Commodity Trading Advisors) trade also on **commodity** and **FX** (foreign exchange) markets. Therefore it would be a good idea to include commodity indices and exchange rates as risk factors, too.

There is also strong evidence that there is predictability in **lagged factors** as shown for example in (Amenc, Bied, & Martellini, 2003). They use a broad multi factor model and then show the utility in using the predictive power of the lagged factors for asset allocation decisions. (Asness, Krail, & Liew, 2001) also find significant exposure to lagged factors; they argue that the beta of a factor and its lagged betas should be summed up to get the real exposure to a certain factor.

(Lhabitant, 2004, S. 215). As many long short equity and relative value hedge funds focus on either value or growth stocks and/or take bets on differences between them it may also be a good idea to include these indices, too.

²³ It has been shown by (Schneeweis & Spurgin, 1999) that the last two factors have explanatory power for hedge fund returns.

²⁴ For details see (Chen, Roll, & Ross, 1986).

²⁵ For details see (Fama, 1981) and (Fama & Schwert, 1977)

Using all these factors in a regression analysis essentially means nothing else but mapping the strategies that are used by hedge funds on a set of liquid and investable securities. And indeed there are attempts to replicate hedge fund index returns with this or similar techniques, the so called hedge fund clones. One of the most prominent proponents is Harry M Kat who designed a product called FundCreator²⁶ and by now all the big investment houses have similar products on offer. But is it really possible on fund level to capture the risks inherent in some fancy hedge fund strategies that act in illiquid and insular markets or to capture the return drivers in special arbitrage strategies with indices built upon market prices of liquid securities? If one looks at the R²s of regressions with these proposed factors, which lie normally around 30%²⁷ the answer is no.

A pragmatic way to solve this problem is to use some aggregate hedge fund returns as factors. As managers assign themselves to certain strategies, if they are listed in a database, it is possible to construct indices for these strategies and for the whole hedge fund universe. Although also the funds within a certain strategy are far away from being a homogenous group, it is likely that on an aggregate level the idiosyncratic risk cancels out and the index provides us with an approximation of the systematic risk inherent in the strategy. Meanwhile there are a lot of differently calculated **hedge fund indices** available from different database vendors and universities and some of these indices are even investable by now. Based on our experience these hedge fund indices better explain the hedge fund returns and we will also include them in our set of regressors. (Keep in mind that we do not want to come up with an economic model for hedge fund returns, we simply want to fill up short track records.) A way to construct a customized index for a single hedge fund will be discussed in section 2.6.4.

2.5.3 Modeling Dependencies between Risk Factors

An interesting question that arises with the use of factor models is how to model the dependencies between the risk factors. The simplest way is to use the historical realizations. But if the historical scenarios are insufficient (because there are simply not enough of them to estimate parameters needed in the optimization reliable), one can think about how to generate new scenarios for the risk factors, which then could be used to generate scenarios for the individual funds.

²⁶ Mr Kat enhanced the initial linear factor approach by using copulas. He explains his views *iter alia* in (Kat, 2007), (Kat & Palaro, 2005); (Kat & Palaro, 2007) and (Kat & Palaro, 2006).

²⁷ According to our experience.

(Lo, 2008) proposes to model the dependencies between the factors based more on experience and on economic intuition than to try to detect them empirically. This may be right, but as we want to focus on quantitative methods we will propose some techniques how to estimate and to model dependencies between a smaller group of variables in a very individual way.

The most “data driven” approach would be some kind of n-dimensional **scatter plot smoothing** or nonparametric density estimation as the fewest assumptions about the structure of the data have to be made. For a nice overview see for example (Scott & Sain, 2005). Although they extend their calculations only to the four dimensional case they argue that, kernel methods are feasible in as many as six dimensions. They do not say anything about the amount of input data needed for the different techniques, but just from looking at the figures they provide it is surely more than we can deliver in our case. As we have a very large number of possible risk factors and only a little amount of data for some of them, these techniques are not appropriate for our needs.²⁸

If one is willing to assume a multivariate normal distribution one can use the **Cholesky decomposition** to generate random outcomes. But as it is not very likely that all the above mentioned factors are normally distributed and as it is even more unlikely that they are multivariate normal distributed, we will consider other techniques as well.

(Morton, Popova, & Popova, 2005) use the so called **NORTA** (normal-to-anything) method of (Cario & Nelson, 1997). The general idea behind this concept is to generate in the first step multivariate normal correlated random numbers via the Cholesky decomposition. Then they set this numbers in the distribution function of a standard univariate normal distribution to get correlated numbers in the interval [0,1]. Then they set these numbers in the inverse distribution function of the marginal distributions of the original data. Mathematically this looks like follows:

1. Generate a multi-variate normal Z with mean zero, correlation matrix Σ_Z , and $\text{Var}Z_i = 1, i = 1, \dots, m$.
2. Form

$$\zeta^i = \begin{pmatrix} F_1^{-1}[\Phi(Z_1)] \\ F_2^{-1}[\Phi(Z_2)] \\ \vdots \\ F_m^{-1}[\Phi(Z_m)] \end{pmatrix},$$

where Φ is the distribution function of a standard univariate normal, and $F_i^{-1}(x) = \inf\{u : F_i(u) \geq x\}, i = 1, \dots, m$.

Figure 10, (Morton, Popova, & Popova, 2005), Formula NORTA Method

²⁸ The amount of input data needed grows exponentially with the number of dimensions.

The problem here is that after this procedure the generated random numbers have another correlation matrix (Σ_{ξ}) than the initial input matrix Σ_Z because of the different marginal distributions (only if all the marginals are normal the correlation matrix would stay the same). So the trick is to find a Σ_Z which leads to a Σ_{ξ} that is identical with the correlation matrix of the original data. This can be done numerically.

We implemented the procedure for the bivariate case in Excel, using the Excel Solver to do the numerical part. The first figure shows the original data (it is the HFR Merger Arbitrage vs MSCI World), the second one 3000 random numbers generated with the NORTA procedure.

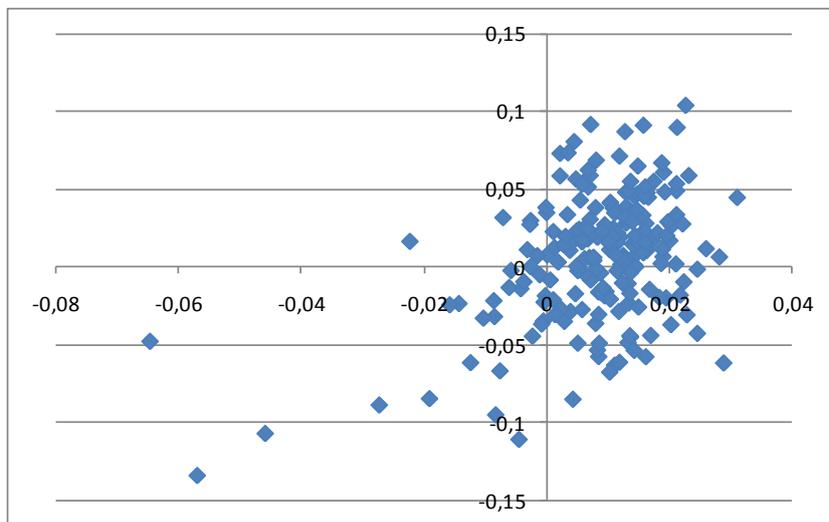


Figure 11, Original Data

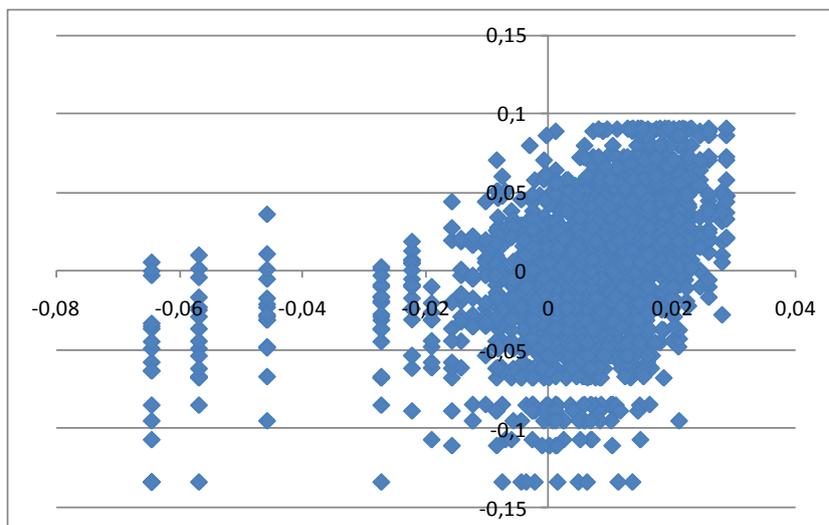


Figure 12, Random Numbers Generated with the NORTA Method

The way this method works can now be seen here very well: as we use the discrete empirical distributions as marginals it generates around each point new points according to the correlation and the marginal distribution along the horizontal and the vertical axis. If we would use some smoothed functions as marginals instead the figure would look a little bit nicer, but the fact that we observe a lot more independent extreme negative outcomes and less co-movement in the extreme negative outcomes than in the original data would remain.

Another way to model dependencies very flexible is to use **copulas**. We used the MATLAB functions `copulafit` and `ksdensity` to fit the copula and the marginal distributions respectively and then generated random outcomes with the `copularnd` MATLAB function to get the following figure:

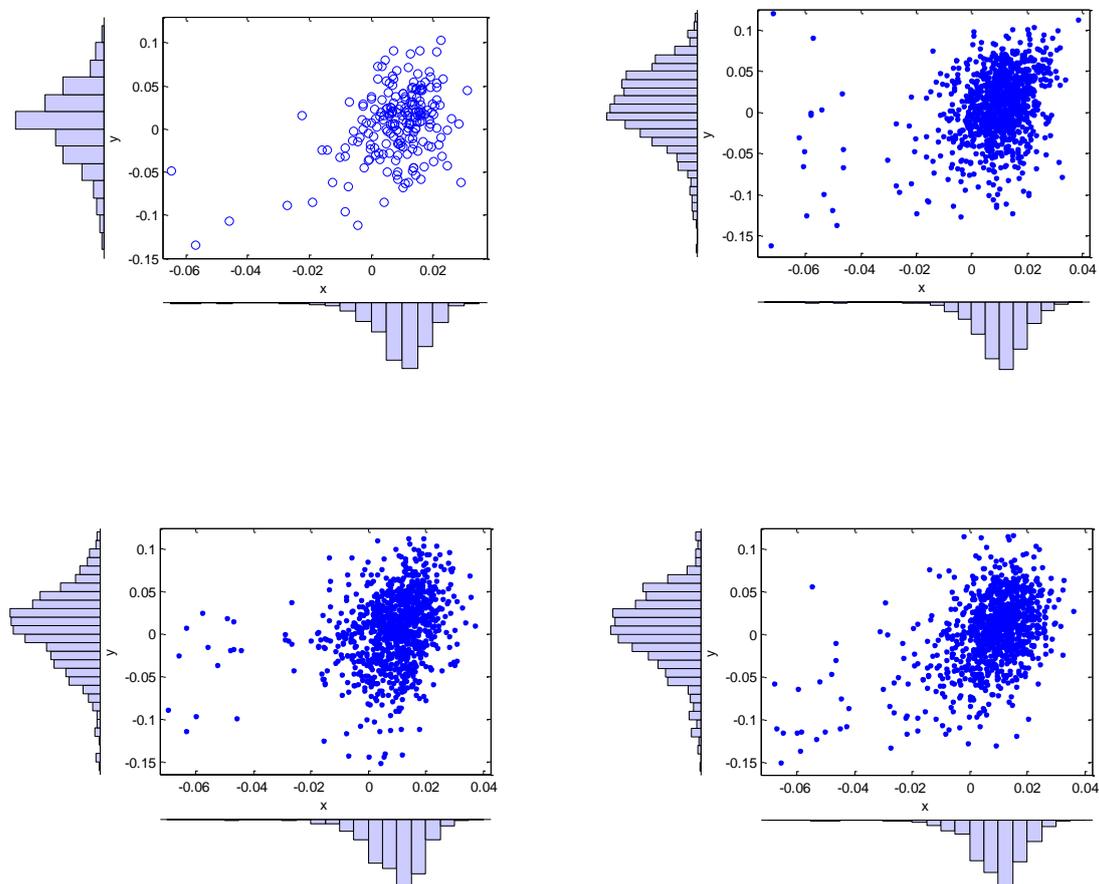


Figure 13, Original Data and Random Numbers Generated with Copulas

The first panel again shows the original data, the other three panels show the random outcome generated using a t-, a Gaussian and a Clayton copula respectively. From this little experiment one can see that the quality of the outcome can differ quite substantially. The problem with this dataset is that there is hardly any dependency under normal market conditions but a strong co-movement in

the extreme negative outcomes. Both the t- and the Gaussian Copula are not able to capture this feature adequately (very much like the NORTA method), because the rho as a measure of the strength of the relationship is mainly determined by the big cloud of independent points. Only the Clayton Copula is able to capture the tail dependency, as it is designed model this effect. To quantify this effect: in the original dataset 2.44% of the observations lie in the -1.8% / -4.8% quadrant. With the NORTA method its 1.2%, with the t- copula it is 1.4%, with the Gaussian copula it is below 1.0% and with the Clayton copula it is 2.3%.

Concluding we can state that it is not so easy to find the right model by hand and it is even harder when trying to do so automatically. And this experiment was just for the bivariate case. Clearly here is a lot room for future research, but for our analysis we will simply use the discrete empirical distribution of the risk factors. This assumes that the future has only and exact these 118 historical states we have observation for, but for the time being this is the best guess we can make.

2.5.4 Non-Linear Dependencies

One issue that is frequently mentioned with factor models in the case of hedge funds is that dependencies between the funds and the underlying risk factors often seem to be non linear because of the use of derivative instruments and/or dynamic trading strategies. There are several ways to deal with this problem:

- Estimate two betas, one for upward markets and one for downward markets. This method is quite illustrative and enlightening for hedge fund indices or funds with a very long track record, but as we will use regression analysis mainly to fill up short track records this approach does not make sense as we may end up with 2 or three data points to estimate the down- market beta.
- Include squared and higher order terms. This is straightforward to implement but not so easy to interpret. On the first view one could think about the squared term as dependency to the risk factors volatility, but in fact this dependency is measured via the coskewness of the fund and the factor. To get an idea about this just insert $(r_m - \bar{r}_m)^2$ instead of $(r_m - \bar{r}_m)$ in the formula for the estimator of the beta. More formally this can be shown via the extension of the CAPM to the so called Higher Moment CAPM. There is also a link to market timing: a manager who shifts successful between market and risk free asset will decrease his exposure on the downside and increase on the upside, leading to a convex function of market

returns.²⁹ Higher Moments seem to be also related to the Fama-French factors as the Fama-French factors become insignificant when higher order terms are added in the regression.³⁰

- Apply option structures on the risk factors. This most frequently used approach in hedge fund literature will be discussed in more detail in the following section 2.5.5.

Other techniques like copulas³¹, smoothing techniques like LOESS³² or phase locking models³³ may work well in a low dimensional case or if the analyst has an idea about the nature of the relationship between fund and factor but they are not appropriate when trying to model dependencies automatically out of a bulk of hundreds of factors.

2.5.5 Option Structure

The **procedure** for the option structure approach is very simple in general: take a risk factor (typically an index), simulate the returns of an option on this index and then calculate the fund return by weighting the index and the option returns. Which options to use will be discussed below. The implicit assumption behind this is to hold the index at a constant level and to buy an option (that expires in one month) at the begin of the month and then execute it at the end of the month, buy a new option and so on, rolling over the whole observation period. The weighting between index and option can be done simply by using both as factors in an OLS regression. So instead of modeling the dependency of fund and risk factor in a complex manner, simply a second return series is generated as an economical reasonable function of the return series of the risk factor. Together the two return series now describe the dependency between risk factor and fund.

The effect (besides the nonlinear dependency) is that the distribution of the generated fund return is no longer normally distributed, a feature frequently observed in reality. The combination of the index and a put on the index for example gives a right skewed distribution. To get a picture of a realistic portfolio we simulated outcomes of three (correlated) indices and of at the money long put options on these indices. The resulting distribution is the left histogram of the figure below. This is a very

²⁹ For details see (Treyner & Mazuy, 1966).

³⁰ For Details see (Chung, Johnson, & Schill, 2006).

³¹ Used for the case of hedge funds for example by (Kat & Palaro, 2005).

³² Used for the case of hedge funds for example by (Favre & Galeano, 2006).

³³ For details see (Lo, 2008, pp. 18-19).

desirable feature for many investors, because it limits the downside risk. The complementary position (a short put) gives a left skewed distribution (right histogram), talking the tail risk for the protective put strategy. All the distributions are constructed (by adjusting the price for the option) to have the same expected return for reasons of comparability.

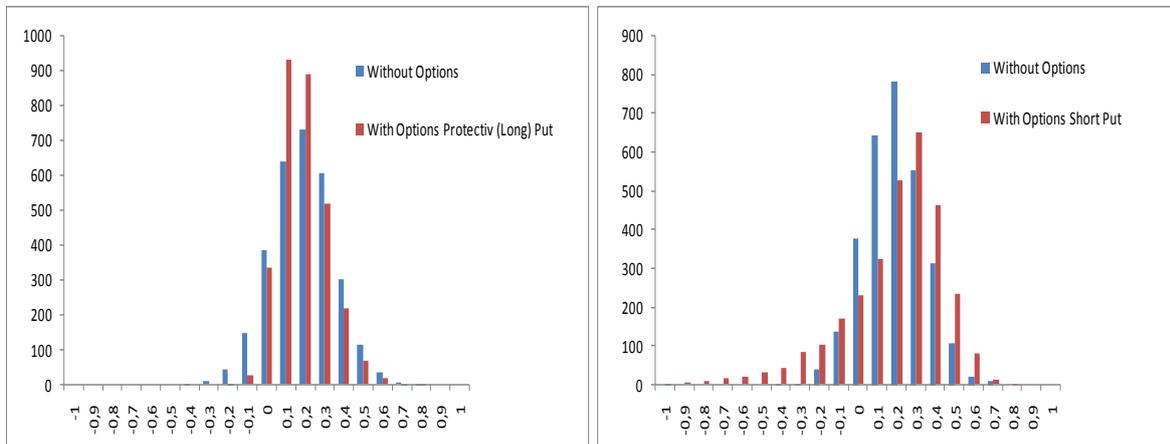


Figure 14, Distributions of Strategies Including Options

A not so trivial question is how to get the price for the option. In principle there are three possibilities: assume fix costs, try to replicate real option prices via the Black Scholes Formula or use real option prices:

The following figure shows the return of a long put versus the underlying asset assuming fixed costs of the option. This provides us with a bilinear function on the underlying. For calculating the fixed cost we propose to assume that the overall return (over the whole observation period) of this option is zero. For a long call this would mean to subtract $\sum_{t=1}^T \max(r_t, 0) / T$ from the payoff each month. By doing so the alpha keeps its interpretation as manager skill in the following sense: The interpretation of manager skill in this context would be that the manager was able to generate an “option like” payoff at lower costs than the fair price with perfect foresight.

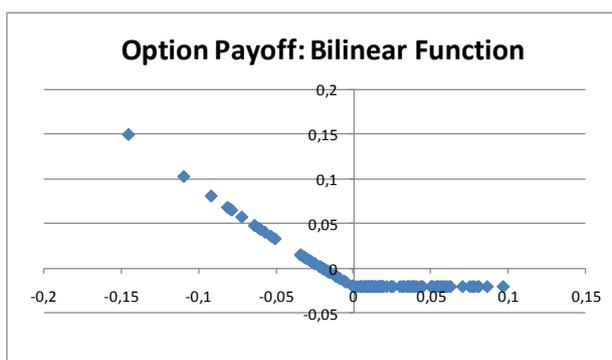


Figure 15, Option Payoff – Fixed Costs

In the next figure the cost of the option is calculated with the Black Scholes formula, taking the historical volatility as input. The index used here is the S&P500 and we used a one year rolling window over the monthly returns to calculate the historical volatility. The typical kink is still observable, although a lot “noisier”. For the Black Scholes formulas see for example (Hull, 2006, S. 361) or (Wilmott, 2001, S. 192). The VBA Code used can be found in appendix 5.1.

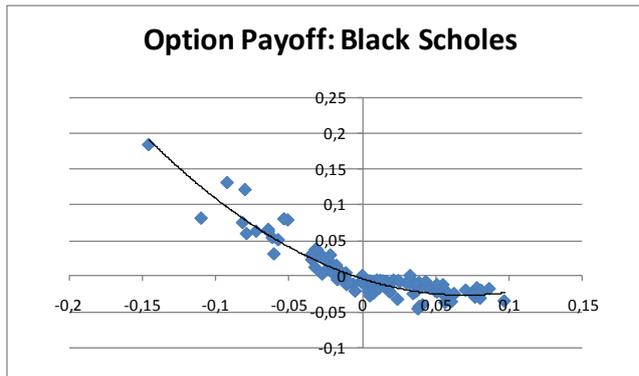


Figure 16, Option Payoff – Black-Scholes Historical Volatility

For the next figure the implied volatility calculated from put options on the S&P 500 as provided by Bloomberg with the Ticker “HIST_PUT_IMP_VOL” was used. We see even more fluctuations and observe that the real prices for the put options are substantially higher than the ones calculated with the historical volatility. The term real price is not totally correct as the implied volatility is averaged from prices of put options with different maturities and strike prices. But, as we do not have access to historical option prices with our Bloomberg account, to use the historical implied volatility is the best way to reconstruct the original option prices.³⁴

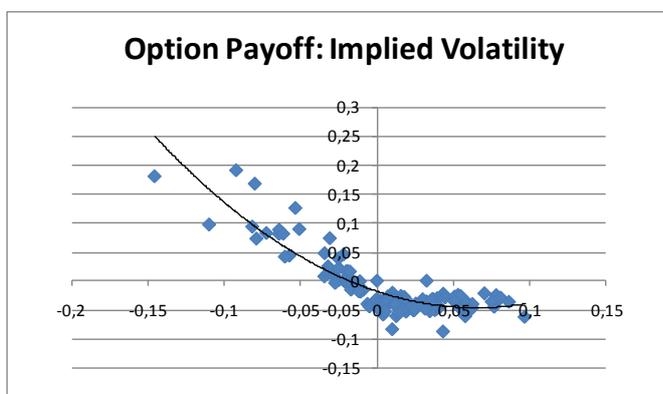


Figure 17, Option Payoff – Black-Scholes Implied Volatilities

³⁴ The implied volatility is backed out from the Black Scholes Formula knowing every input factor (including the observed price of the option) but the volatility. By inserting the implied volatility in the Black Scholes formula one can reconstruct the price the implied volatility was calculated with.

The last figure shows the return of a 3 month 5% out of the money put option on the DAX. It is backed out from a really executable option strategy that is represented by the DAXplus Protective Put Index.³⁵ Here the option payoff characteristics are quite hard to observe, especially the high prices for the put in volatile markets make the relationship more linear.

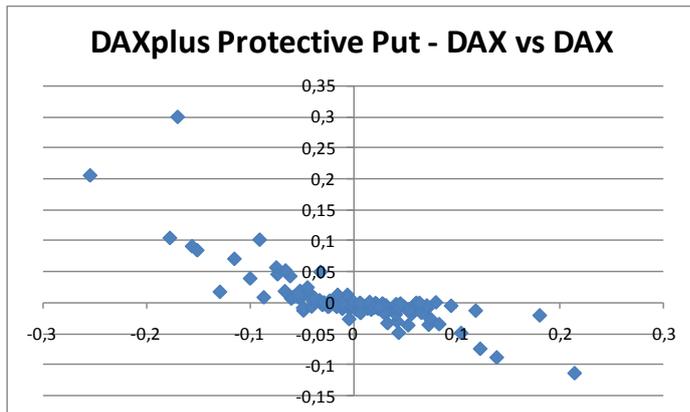


Figure 18, Option Payoff – DAXplus Protective Put

We decided to use the **bilinear function** with assumed constant costs. We did so because it turned out to be extremely difficult to replicate real option returns with just having information about the index returns, the implied volatility and the interest rate.³⁶ In addition we made the experience after using really executable option strategies like the DAXplus and EUROSTOXX option strategy indices for over one year, that they hardly ever were able to explain a funds return.³⁷ This may not come as a surprise as it is really very unlikely that a manager uses exact the strategy represented by a particular option strategy index, using options with exactly the same strike price and the same maturity. On the other hand the use of the bilinear functions sometime was able to push the R^2 in a meaningful way quite substantially.

There are also several other **advantages** of the “option – bilinear function” approach: It can easily be extended to the multivariate case. It is not dependant of a single price for a specific option but generates “option-like” returns. The interpretation of manager skill as the ability to create “option-like” returns at lower costs than the fair price with perfect foresight is not implausible. In contrast to

³⁵ See (<http://deutsche-boerse.com/>)

dbag/dispatch/de/binary/gdb_content_pool/imported_files/public_files/10_downloads/50_informations_services/30_Indices_Index_Licensing/40_downloads/50_Strategy_Indices/M_FL_PP_e.pdf; The calculation of this index includes funding costs and takes real prices of one specific option.

³⁶ Unfortunately a Bloomberg professional account is necessary to get historical option prices.

³⁷ For the S&P 500 and the Nikkei such indices are not available. Nonetheless we constructed, based on the methodology of the DAXplus und EUROSTOXX indices, our own indices. But they also did not provide meaningful results.

other techniques to model nonlinear dependencies the analyst keeps an economic intuition as dynamics of options are well known; it is easy to implement via OLS; it is very flexible via the combination of different option types, setting different moneyness and different weights on the risk factor and the option respectively. To demonstrate this flexibility the following figures are provided:

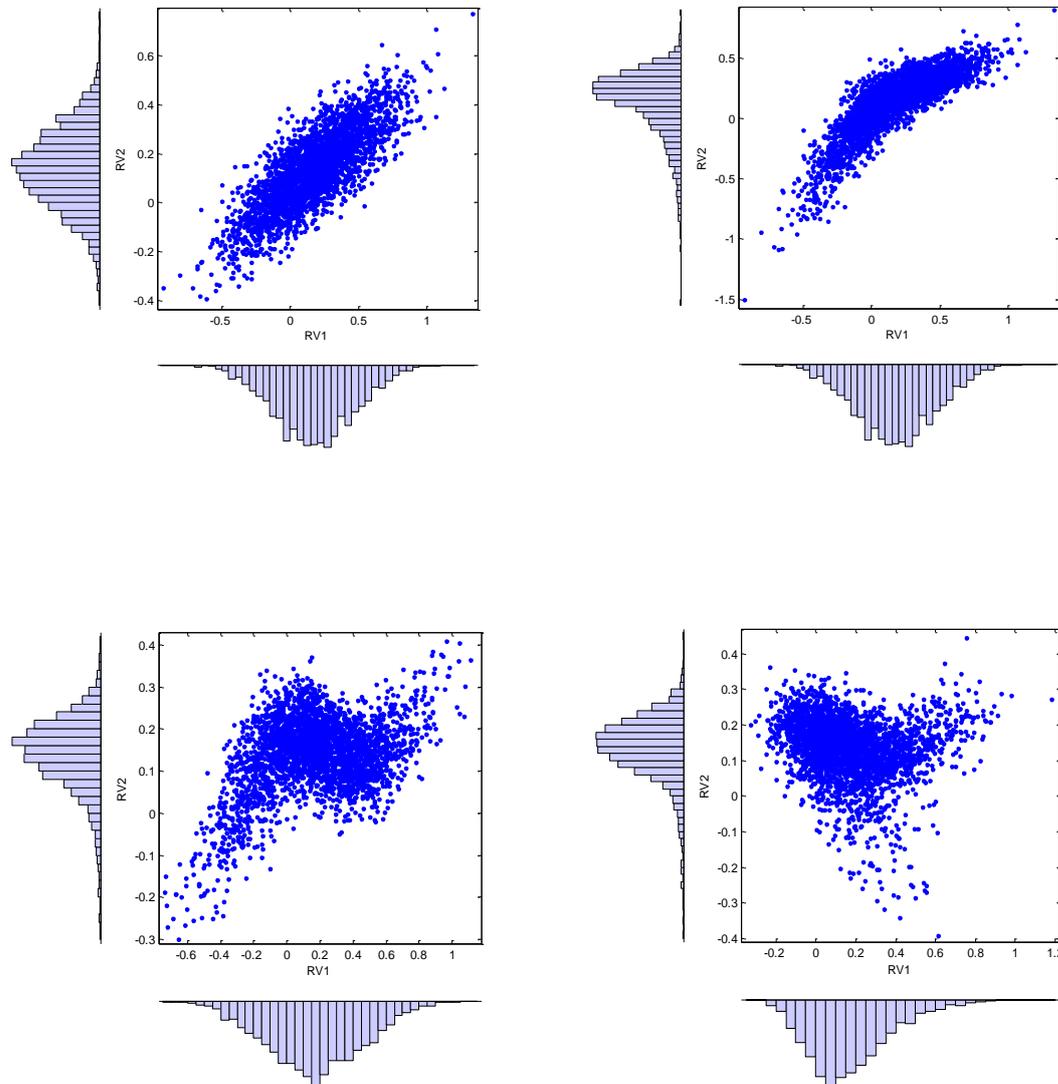


Figure 19, Dependency Structures Created with Option Payoffs

The figures were generated as follows: The assumption is that we have two funds that invest both in three indices and in options on these indices. The indices are multivariate normal distributed. So if the funds only invest in the indices the bivariate distribution of their returns will be multivariate normal, too. This can be seen in the upper left panel of the figure above. If one of the funds also invests in options, namely in short puts with strike price zero, it looks like the upper right panel. As we set the expected return of the options to zero (via adjusting the constant price of the option) the

expected return for RV2 is the same in figure one as in figure two. In the lower left panel investor two holds long puts with strike price 0.4 in addition. This panel is quite illustrative as the marginal distributions are (more or less) normal, but the dependency structure is quite unusual. Just using the standard dependency measure beta, which is linear, would lead to outcomes similar to figure one. In the lower right panel investor one invests in long puts with strike zero in addition.

The second set of figures should compare the option structure approach to the copula technique. Therefore we plot the “copula” that would be required to generate the above dependency structures when assuming normal marginal distributions. This can be done by standardizing the variables and then map them on the [0,1] interval by computing the cumulative probabilities of the standard normal distribution for the standardized returns.

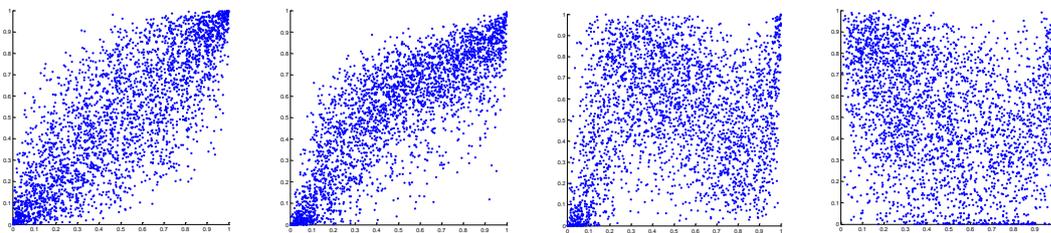


Figure 20, “Copulas” of Dependencies Created with Option Structures

The first panel really is a copula, a Gaussian one, generating – together with the normal distributed marginals – the multivariate normal distribution of the upper left panel in Figure 19. The next panels show that quite unusual copulas would be necessary for the dependency structures in Figure 19. To my knowledge there is no copula that even looks similar to panel three. To be precise the last three panels cannot be copulas as they are not uniform distributed, a feature required for a copula.³⁸ But this problem is not hard to circumvent in the end, one simply has to assume the empirical distributions as marginals. The difference is shown for figure four:

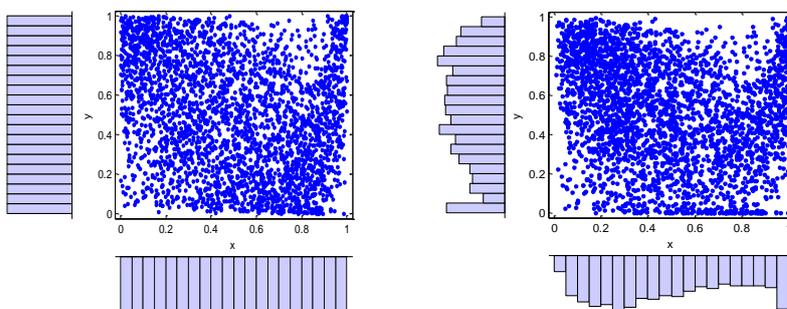


Figure 21, “Copulas” with Empirical Marginals vs “Copula” with Normally Distributed Marginals

³⁸ For an exact definition of a copula see for example (McNeil, Frey, & Embrechts, 2005, S. 185)

Besides the fact that the option structure approach is very flexible and allows us to keep an economic intuition, there are some interesting **findings** on the option structure approach in the literature. The central one is that most hedge fund strategies are similar to writing a put on a broad, traditional index.³⁹ A right-skewed distribution like produced by the protective put strategy like exhibited in the left panel of Figure 14 is very desirable for most investors and therefore relatively expensive. Hedge funds seem to take the counter position here: they are willing to take downside risk and therefore get a premium under normal market conditions – at the end this is a provision of insurance against extreme market events.

This may explain the high turnover and **short lifetimes** in the hedge fund industry: the fund gets its premium for taking downside risk and generates impressive excessive return.⁴⁰ With such an excellent track record the company markets the fund heavily and earns high fees for the high alpha they seem to be able to generate. When a significant negative market event takes place the fund realizes a huge loss and gets liquidated.⁴¹ The company behind creates a new fund, and starts over again. Clearly this is a very critical view on the hedge fund industry but at least for part of the industry there surely is some truth in it.

Another frequently cited finding is the one of (Fung & Hsieh, 2001). They use lookback straddles to explain the return of funds following “trend following” strategies. Both, returns from (lookback) straddles⁴² and trend following strategies have strong positive skewness and positive returns during extreme up and down moves in the markets. And as the payout profile of a perfect market timer is like a portfolio of bills and a call option⁴³ we propose to use a put, a call and a straddle profile to model the hedge fund returns.

³⁹ See for example (Agarwal & Naik, 2004)

⁴⁰ The simplest example for such a strategy is to write out of the money puts on the S&P 500. (Lo, 2008, p. 7) calculates the returns of this strategy for the timeframe January 1993 to December 1999 and compares it to the returns of holding the index: 3.6% monthly mean return versus 1.4%, 6 negative months versus 36.

⁴¹ -18,3 and -16,2% in August and September 1998 in the case of the (Lo, 2008, p. 7) example.

⁴² Empirically it does not make much difference if lookback straddles or normal straddles are used to explain the fund's returns. The lookback straddle has the additional feature of path dependency. Confer (Fung & Hsieh, 2001).

⁴³ Confer (Merton, 1981).

2.5.6 Principal Component- and Factor Analysis

The methods described above rely not only on the quantitative data given but also on some assumptions about the economic structure behind, i.e. there are certain risk factors, chosen by the researcher, that determine the return of a fund. Hedge fund indices go away from this assumption, using only the data given by the funds themselves. We argued that via the aggregation we may be provided with systematic risk factors inherent in the hedge fund world. But there are other techniques - than more or less simply taking (weighted) averages - that are able to extract factors from the data, like the factor analysis and the principal component analysis.

The first one we want to discuss is the principal component analysis (PCA). PCA is a method to detect the (unknown) factor structure that is hidden in a data set and to reduce the number of variables. **The idea behind** the PCA is to find a new set of coordinates (one axis per variable) in a way that information is best reflected. The first coordinates or “principal component is the linear combination of the original variables that accounts for as much variation in the data as possible. Each succeeding principal component is the new linear combination of the original variables that accounts for as much of the remaining unexplained variation as possible” (Lhabitant, 2004, S. 186) and stands orthogonal on (=is independent of) the preceding principal components.⁴⁴

As the PCA is frequently used to **visualize** high dimensional data in a low dimensional space the following sentence gives a nice interpretation: “If a multivariate dataset is visualized as a set of coordinates in a high-dimensional data space (1 axis per variable), [the first two principal components supply] the user with a 2D picture, a shadow of this object when viewed from its most informative viewpoint.”⁴⁵

One idea how to use the PCA with hedge fund data is to perform the PCA to find factors that then can be used to **explain** the **fund’s returns**. If we are lucky the principal components reflect the risks inherent in the hedge fund industry, not only the ones that could be modeled with available indices, too, but also the hidden ones like liquidity risks.

We performed a little **experiment** to check whether the PCA is able to find powerful factors on fund level or not. We use the ‘only USD’ HFI database for this purpose. The analysis is performed in MATLAB using the princomp function.

⁴⁴ For a mathematical precise discussion of PCA see (Fahrmeir, Hamerle, & Tutz, 1996, S. 661-677).

⁴⁵ See (http://en.wikipedia.org/wiki/Principal_component_analysis).

In the first turn we use only funds that have entries for the whole observation period (Jun 1998 to Mar 2008). This leaves us with a sample of 336 funds having each 118 monthly returns. For robustness reasons the analysis was repeated with funds that have entries for the last 36 months (sample size: 1650) and with standardized data.⁴⁶ The results are similar. The following figure shows the marginal (bars) and cumulative (line) variance explained by the first 20 principal components and was generated using the MATLAB pareto function:

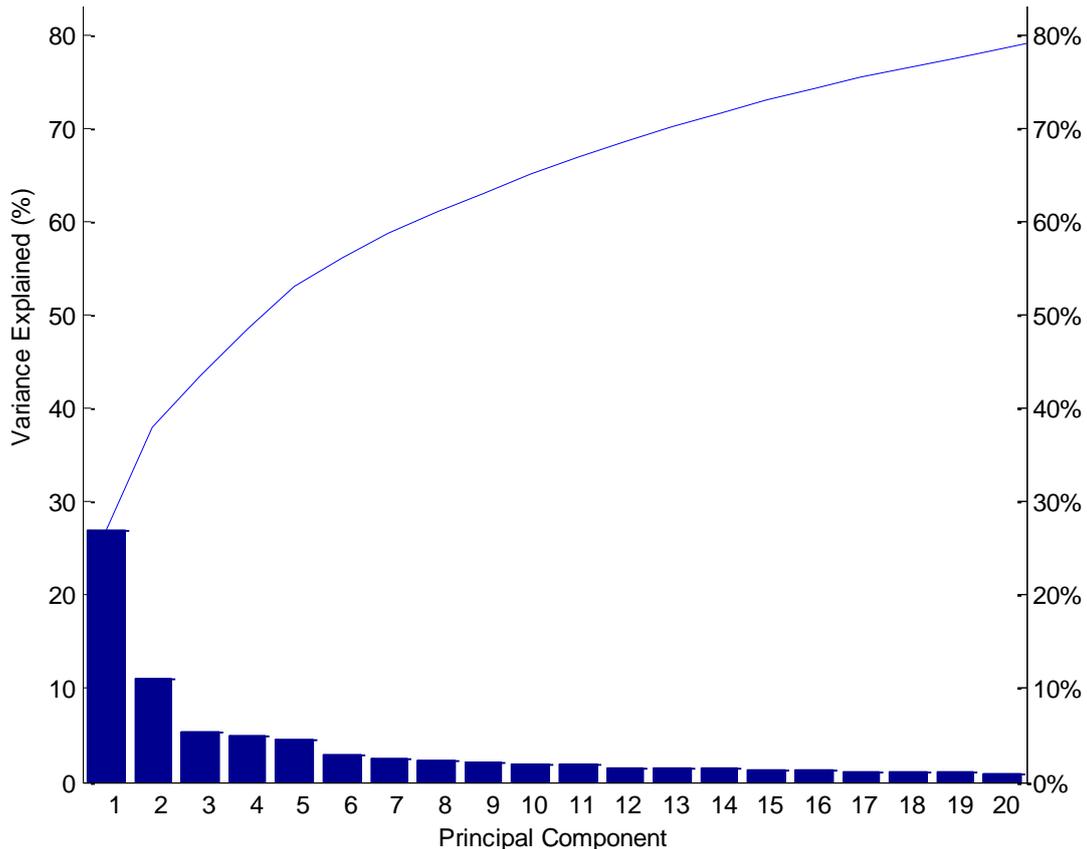


Figure 22, Pareto Plot of Variance Explained of the First 20 Principal Components

There seem to be at least two relevant factors. To test if the PCs are significant in a statistical sense we follow the approach by (Johnstone, 2001) who calculates the distribution of the eigenvalues under the assumption of a purely random correlation matrix. To be on the safe side one would only accept PCs with a corresponding eigenvalue higher than the maximal eigenvalue of the random matrix. In our case this would mean that the first seven PCs are non random, although PC 6 and 7 are very close to the limit.

⁴⁶ To use standardized data mitigates an effect not always intended: the PCA interprets high variance as high level of information.

To find out if the principal components have an economic interpretation we try to find indices that are highly correlated with them. The first principal component (PC) seems to capture the broad movements of the hedge fund universe, an aggregate factor that is most highly correlated to the CSFB Multi Strategy Index (0.956) and the EurekaHedge Fund of Funds Index (0.948). The second PC is more or less uncorrelated with all available indices, as well as fourth and fifth PC. Only the third one has a significant correlation, namely 0.804, to the bund future. But as it is unlikely that the bund future reflects a driving risk factor this more likely is a statistical artifact. With standardized data the second PC was highly correlated to the MSCI Directional Trading (0,937) and EuroHedge Managed Futures USD Index (also 0,937) but on the other hand the first two PCs were able to explain only slightly above 30% of the variance.

There could be done a lot of further analysis like examining which kinds of funds (strategy) contribute mainly to which PC. (Fung & Hsieh, 1997) did so their "classical" paper from 1997 (indeed it is the most heavily cited paper in the literature about hedge funds) and they really were able to assign meaningful names to the first five PCs by checking the strategy self description of the funds that are highly correlated with the PCs. But as this their PC did only account for 12% of the variance and the first five ones together only for 43% it can be doubted that there really is much information in them. (Christiansen, Madsen, & Christiansen, 2003) did a similar analysis six years later and their first five PCs explained more than 60%. Our explanation for this effect is that this is a sign for consolidation in the hedge fund industry as more funds seem to be highly correlated among each other letting the PCA find more powerful PCs. High correlation indicates that these funds follow similar common strategies, no longer offering a unique return profile.

Another way how to determine the usefulness of the PCA can be found in (Alexander & Dimitriu, 2005). They calculated minimum variance portfolios with the best funds in terms of alpha. For the calculation of the alpha they used different factor models, one of which had the PCs as factors. They conclude that the alpha ranked minimum variance portfolios based on PC factors are best performing out of sample in terms of information ratio and turnover compared to factor models that have an economic interpretation.

For our study we conclude that the PCA is able to extract one factor that reflects quite well the overall movement of a sample, but that the other PCs seem to be statistical artifacts rather than real risk drivers. Clearly the PCs that are uncorrelated to the known risk factors and hedge fund indices could reflect some unknown, hidden factor but as a lot of research is done on risk factors it is quite unlikely that the PCs two to five all capture unknown factors.

There are several reasons why the **PCA** can **fail**:

- The PCA assumes linearity between the PCs and the data. As argued above linearity between risk factors and fund returns must not hold true for hedge funds. To use a kernel PCA (Schoelkopf, Smola, & Mueller, 1998), which would solve this problem, would be an interesting research project.
- PCA only works optimal “under the Gaussian assumption. For non-Gaussian or multi-modal Gaussian data, PCA simply de-correlates the axes.”⁴⁷ Hedge fund returns are often not normally distributed.
- PCA assumes that large variances reflect important dynamics. By column wise standardization we accounted for this problem.

But there are better ways to use PCA than trying to extract hidden risk factors out of a sample of individual hedge funds. One example is to apply the PCA not on fund but on index level. (Amenc & Martellini, 2002) did such an analysis that was the basis for a new set of hedge fund **indices**, the EDHEC alternative indices.⁴⁸ They argue that the databases (out of which the available hedge fund indices are calculated) each account only for a fraction of the hedge fund universe and hence the indices do not reflect the whole universe adequately. But instead computing an equal or value weighted “hyper”index they suggest to weight in such a way that the largest possible fraction of information is captured, the “best possible one-dimensional summary”. This is nothing else but the first PC.

Together with our finding that the first PC on fund level reflected some overall movement we recommend to use the PCA to extract a one dimensional summary, too, but in a different way. As the main aim of this section is to fill up track records we developed the following idea: for the fund you want to fill up find the n-most similar funds that have a longer track record and then extract the best one dimensional summary out of their returns. How to come up with the n-most similar funds will be discussed in section 2.6.

Like the PCA the **Factor Analysis** is a data reduction or structure detection method. The main difference is that with factor analysis orthogonality is not required.⁴⁹ In some cases this makes it easier to interpret the factors economically. The problem we faced with factor analysis was that the

⁴⁷ http://en.wikipedia.org/wiki/Principal_components_analysis

⁴⁸ See (<http://www.edhec-risk.com/>).

⁴⁹ In a mathematical more correct sense PCA and factor analysis are two very different approaches: PCA is an analysis of the eigenvalues of the covariance matrix and factor analysis uses a maximum likelihood estimator.

factoran MATLAB function requires a positive definite covariance matrix of the data. As we have more variables (336) than observations (118) the covariance matrix is not guaranteed to be positive definite and it definitely was not for our sample.

2.5.7 Model Selection

The big question that remains is how to come up with a meaningful factor model out of the huge amount of possible factors we discussed. In all papers that concern factor models the authors restrict themselves to a limited, carefully selected set of risk factors. What we have done in the previous sections was nothing else but to collect and order the risk factors discussed in the different papers. At the end we are left with more than 100 risk factors to choose, in the Benchmark database there are 270 ones, including several sets of hedge fund indices. If we want to account for non linearity the number of possible factors grows dramatically on top of this: Clearly it does not make too much sense to assume nonlinear relationships to hedge fund indices, but there are still more than 100 equity, currency, commodity and fixed income indices where nonlinear dependencies are possible. To add squared returns doubles the number of possible regressors, as well as with the use of lagged factors. When adding option payoffs the number gets multiplied by 9 if we use three types of payoff profiles (put, call and straddle) and three different strike prices.

We did some first experiments to find out what procedure is able to come up with a meaningful result. The design was as follows: we took a set of funds that all had return data for a certain time window (say 48 months).⁵⁰ Then we took one fund out of those and used only a proportion of the data to fit the model (say 12 months) and then checked the quality of the model with the remaining data. To use all possible factors as regressor clearly would not be sensible, we used the MATLAB `stepwisefit` function to select the appropriate regressors instead.

The experiments with the option structure approach showed very unsatisfactory outcomes: In sample there was nearly perfect fit, out of sample the results did not make sense at all. This was mainly due to very high coefficients for the option payoffs. Although we were able to produce good and meaningful results with the option structure approach when having an idea about the funds risk exposures and specifying the model by hand we were not able to find an automatic model selection procedure that can do so as well by now. Also with the squared returns and the lagged factors we

⁵⁰ For robustness reasons we did not use only the last 48 months but a 48 months period randomly chosen over the whole observation period.

had similar difficulties. It seems that it is necessary to feed some kind of prior information about what the fund is doing and on which markets it acts to get practicable results.

At the end we did find that to use all indices (hedge funds and traditional) as possible regressors and to set very tight parameters⁵¹ for the stepwisefit command produced the best results. Clearly a stepwise fitting is very likely to stop at a local minimum, but this is what we want it to do in sake of robustness, because the perfect in sample fit would lead to meaningless results out of sample. The MATLAB code can be found in appendix 5.4.

2.6 Cluster Analysis and Qualitative Data

2.6.1 Introduction

In the section above some techniques that can be used to fill up the track record of a hedge fund based on factor models were discussed. The remaining problem is, if the fund for example has two years of track record, we have only 24 data points to estimate the factor model as hedge funds report their return only on a monthly basis. Another problem that arises is that the factor model representations are often not stable over time as many funds, especially global macro and multi strategy, act very opportunistically across the different markets. The easiest way to see this is to look at rolling correlations. A nice example here is the dependency of the CS/Tremont Multi-Strategy Index versus the S&P 500: In the early 2000s the correlation was negative but then with the enormous inflows of money into multi strategy funds it became more and more difficult for managers to generate uncorrelated returns and the correlation of the index became positive (around +0.5). Another explanation for the augmented correlation can be that multi strategy funds opportunistically increased their exposure to the rising stock markets. Evidence for this explanation comes from the fact that correlations are now very low again.

So fitting a model with a relatively short time window and then assuming stable parameters for the past may not be the best approach. An idea to overcome this problem could be to look for funds that are similar to the one considered but that have a longer track record. Clearly this assumes that fund managers that fall into one cluster always act in a similar way and shift their exposures more or less simultaneous, but this assumption is more realistic than stating stable risk factor exposures over time.

⁵¹ Include only factors with a p-value of 0.01 and kick them out with a p-value of 0.025.

The easiest way to incorporate this idea is to use hedge funds with a longer track record as regressors in a **univariate regression** analysis and to take the two or three best fitting ones to calculate an “individual index” for the fund. Clearly the betas should be positive and around one as it is not possible to go short on hedge funds and a different leverage level stable over a longer time period is very unlikely, too. A multivariate analysis would also not make too much sense as a factor model representation with hedge funds as factors is even more unlikely to be stable than with indices reflecting certain risks. The question why not to simply use readymade hedge fund strategy indices can be answered by the fact that funds within a certain strategy still are a very inhomogeneous group and that such an index may be not specific enough to capture the behavior of a certain fund.

The problem that remains also with univariate regression on hedge funds is that we have only a very limited number of data points to estimate the coefficients. Fortunately there is not only return series available in most databases but also **qualitative information** about the hedge funds that can be used to pool the funds into smaller and more homogeneous groups.

In the **HFI database** we have information about the strategy self classification, the assets under management, the administrator, the prime broker, the CEO, the marketing manager, the main office location, the domicile, management- and performance fee, hurdle rate and high water mark, minimum investment, whether the fund is still open to investment or not, the legal advisors, subscription and redemption terms, the auditor, lockup period, currency and data vendor name.

The **Benchmark database** has similar entries and in addition some interesting items about the assessment of the hedge fund analysts about the funds risk exposures. There are also text files available – a detailed strategy description and a short summary. Unfortunately we will not be able to use this text information at this stage, but the in house classification according to market risk, trading strategy, status (is the fund investable for Benchmark), regional exposure and the assigned peer index will be used to group similar funds together.

Besides the data available in commercial databases there might be other qualitative information relevant for risk and return. One can think of the ownership structure of a firm, incentive structures for the managers, if the key managers invest considerable amount of own money in the fund and so on. There are quite long due diligence questionnaires that try to find out more about the funds organizational stability, its IT infrastructure, the investment philosophy, the investment process, the

accounting and reporting and so on. Even things like marital status, the car driven by the managers and other **private information** is considered by large fund of hedge funds that invest so much money that the managers are willing to answer such question. Information that can be used also by a smaller fund of funds is if a university endowment is invested in a fund or not. Normally this is a good sign and funds advertise this.

The question if the qualitative data is considered to be useful or not was not discussed yet. It is not unlikely that this information has, as some kind of idiosyncratic risk, no influence on the returns. At least for characteristics like fund **size** and fund **age** there exist several studies that show their significant influence on the fund's returns. (Lhabitant, 2004, S. 195) for example states for fund size: "For bond arbitrage and market neutral performance (Sharpe ratio) degrades rapidly as fund size is increased, while this effect is not as dramatic or reversed for other strategies." And for fund age: Younger funds have a significant better Sharpe ratio, "perhaps due to size and nimbleness [... But] these results should be taken with caution, because younger funds are likely to be subject to a greater reporting bias."

(Zhong, 2008) finds that "fund-level flow has a positive (negative) impact on a fund's future performance for smaller (larger) funds, while strategy-level flow (flow into the strategy to which a fund belongs) always has a negative impact on the fund's future performance." Unfortunately we do not have a time series on the assets under management available.

(Favre & Galeano, 2006) find a significant albeit weak relationship (R^2 0.035) between hedge fund returns and the **redemption delay**: The longer the delay, the higher the return (up to ten days, then the sign shifts). This is plausible insofar as redemption delay may be a proxy for illiquidity exposure and the higher return an illiquidity premium.

2.6.2 Missing Data and (Dis-) Similarity for Qualitative Data

There were two problems with the Benchmark database: it is "written" in Oracle and not all fields are really filled for all funds as the funds are in different stages of the due diligence process. Therefore in a first step the oracle data get imported and ordered into a .mat file. In the next step the non-numeric data gets transformed into numeric one (e.g. the entry in the field "strategy": "event driven" was transformed into 302, where the 3 indicates the category "strategy" and the 02 the attribute "event driven").

Next the vector, that contains the **numerically coded** qualitative information (e.g. [104 201 ...]), gets transformed into a binary vector. This works as follows: The initial binary vector contains only zeros and has the length (number of different attributes for category 1) + (number of different attributes for category 2) + Then the vector gets filled with ones according to the attributes of the specific fund. The number of different attributes for category 1 (“currency”) is 4: USD, EUR, JPY and ZAR. If the currency of the fund is EUR the first four entries of its binary vector are [0 1 0 0]. The next category is “strategy” with nine different attributes. If the fund has the second strategy the first 13 entries of the binary vector are [0 1 0 0 0 1 0 0 0 0 0 0]. If an attribute is missing all entries for the category stay zero.

As not all information is equally important, this vector gets multiplied with a **weights vector** of the same length. If currency is not so important and strategy is very important, this weight vector could look as follows [0.2 0.2 0.2 0.2 10 10 10 10 10 10 10 10 10 ...].

With this vector one can already compute a similarity measure. As we have categorical or nominal variables only two measures make sense: The Simple Matching Coefficient (SMC) and the Jaccard coefficient (JC).⁵² The first one is calculated by dividing the number of matches by the number of attributes. Clearly this does not make too much sense in our case as there will be a lot of 0-0 matches by construction. Therefore it will be better to use the JC. Here only the 1-1 matches are in the numerator and it is divided by the sum of all entries were a one occurs (0-1, 1-0, 1-1). A JC similarity matrix can be calculated in MATLAB easily with the pdist and the squareform function.⁵³ In Appendix 5.3 the qualiMJ output gives the results.

But there occurs a problem when calculating the JC in the above described way: Assume two funds that have only one qualitative information available and that they match in this category. Their JC is then $1/1 = 1$. Now assume two other funds that have three qualitative information available and match in one category. Their JC would be $1/(2+2+1) = 0.2$. Clearly this does not make too much sense as the probability that the first two funds match in the two unknown categories is not 1.

To circumvent this problem one could simply divide the matches by the number of categories,⁵⁴ or in the case with the weights, by the sum of the weights. The output qualiMM gives the results. But this

⁵² See (Tan, Steinbach, & Kumar, 2005, S. 65-83) and (Fahrmeir, Hamerle, & Tutz, 1996, S. 440-452) for an overview of proximity measures.

⁵³ The pdist function calculates the JC as a dissimilarity measure. To get the JC like described above one simply has to calculate $1 - JC(pdist)$.

⁵⁴ This would be nothing else but the SMC on category level.

methods is radical in the other way round, it assumes that the unknown categories do not match for sure.

From our point of view the best way is to take into account the **probability** that a certain attribute occurs. This can be done by adding the probability⁵⁵ that a fund has the same attribute in a category (for which it has no entry) like the fund it get compared to (and which has an entry for this category) to the number of matches in the numerator. The categories where there are no entries for both funds still count as mismatch. The results get stored in the qualiMP matrix.

The **drawback** of this method is that it does not take into account that some attributes within a category are more similar than other ones, only an exact match (or the probability for an exact match) counts. But in the category “trading strategy” “L/S Equity Sector” and “L/S Equity Market Cap” are clearly closer than “L/S Equity Sector” and “Fixed Income Arbitrage”. For future development one can think about defining distances between attributes within one category.

With the qualiMP proximity matrix one could already run a cluster analysis or find the n-most similar funds for writing back a track record. But maybe it is more advisable that quantitative data is taken into account, too. One could think of many ways how to do this, for example doing a qualitative pre screening and then do the analysis solely based on quantitative data, or the other way round. One could calculate a proximity matrix based on quantitative criteria and one based on qualitative criteria and then get an overall matrix as a weighted sum of the two matrixes. This sounds a little bit easier than it is in the end, as different proximity measures may not be addable in a meaningful way, and clearly to sum up does not take into account interactions between qualitative and quantitative data. For the purpose of finding the n most similar funds for filling up a track record a ranking system with minimum limits of similarity may be the best idea. So there is a lot room for future development, for the time being we want to present some little experiments with clustering according to certain risk measures.

2.6.3 Cluster Analysis with Return Characteristics

In this section we present some outcome of experiments with cluster analysis among hedge funds along certain **return characteristics**. As described above one could try to integrate qualitative data by getting a weighted sum of the proximity matrixes, but here we want to limit ourselves to quantitative data.

⁵⁵ Weighted according to the “importance vector”.

As this experiment has mainly illustrative character we limited the analysis also to funds that have entries for the whole observation period, out of the funds in the 'only USD' HFI database. This leaves us with 336 funds to work with. The following measures were calculated: expected return, volatility, worst month, (empirical) Value at Risk, (empirical) Conditional Value at Risk, correlation to the MSCI World Equity and the mean performance during months in which the MSCI World was negative.

With the measures for all funds one can easily calculate a **proximity matrix** in MATLAB with the `pdist` function. This function provides a lot of different metrics to calculate a distance and the choice of the right metric is crucial for the quality of the outcome. See appendix 5.5 for descriptions of the metrics available.

For the purpose of finding the n-most similar funds for filling up the track record of one specific fund this proximity matrix is sufficient, as the n would be quite small in this case and the proximity of the other funds among each other does not matter. But if one wants to create bigger peer groups, or divide the whole hedge fund universe in clusters one has to group funds together according to certain rules. In classical cluster analysis there are two main different ways how to cluster: hierarchical and partitional methods.

With the **hierarchical clustering** the most common methods are the agglomerative ones in contrast to the divisive ones. The idea with agglomerative hierarchical clustering is to progressively merge funds to bigger and bigger clusters. The following figures show in a simple example how this works in the two dimensional space:

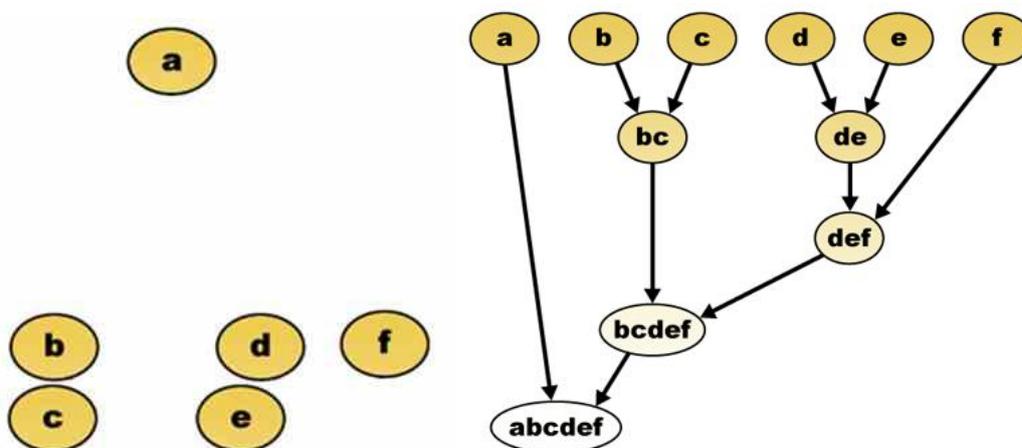


Figure 23, (http://en.wikipedia.org/wiki/Cluster_Analysis), Hierarchical Agglomerative Clustering

Always the most similar objects were merged. These objects can be funds or clusters. Among two funds it is easy to determine the most similar one (just look up the proximity matrix) but it is not so easy to say how similar two clusters are. There are several rules how to determine this and the MATLAB linkage function provides a number of such rules, like complete (uses the furthest distance between two funds in the two groups – for the example above the distance between the bc and the def cluster would be represented by the distance between f and c), single (the shortest distance; c and e in our example) or average (the average distance between all pairs). The rules seem to be simple, but to choose the wrong one could lead to very dissatisfactory results.

There are some theoretical guidelines what metric and what linkage rule to use⁵⁶ but in the end one has to play around with the data to see which combination of metric and linkage rule makes sense. The decision whether to use standardized data or not depends on whether a larger scale corresponds with high information value or not. In our case we found that the Euklidian distance together with the complete linkage and un-standardized input data brought the best results. With the dendrogram function the outcome can be visualized in a tree similar to the one above.

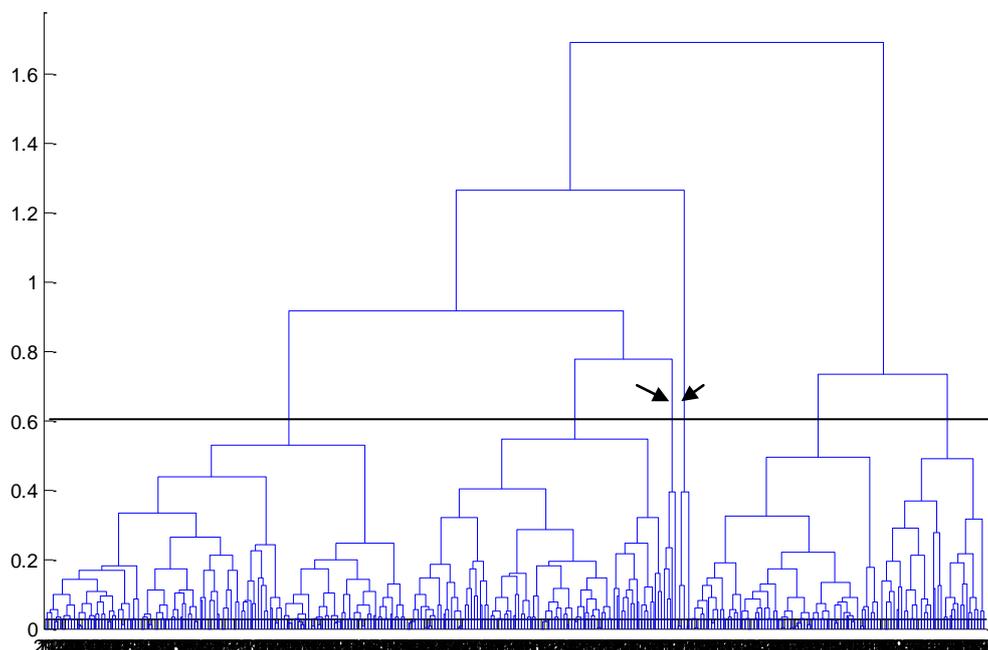


Figure 24, Dendrogram Clustering with Return Characteristics

If the tree gets cut at 0.6 (Y-axis), which can be done with the cluster command, we get 6 clusters. The two clusters marked with arrows are a little small, but indeed they have very special characteristics, as can be seen in the following figure:

⁵⁶ Most of the more complicated rules only work with the Euklidian distance.

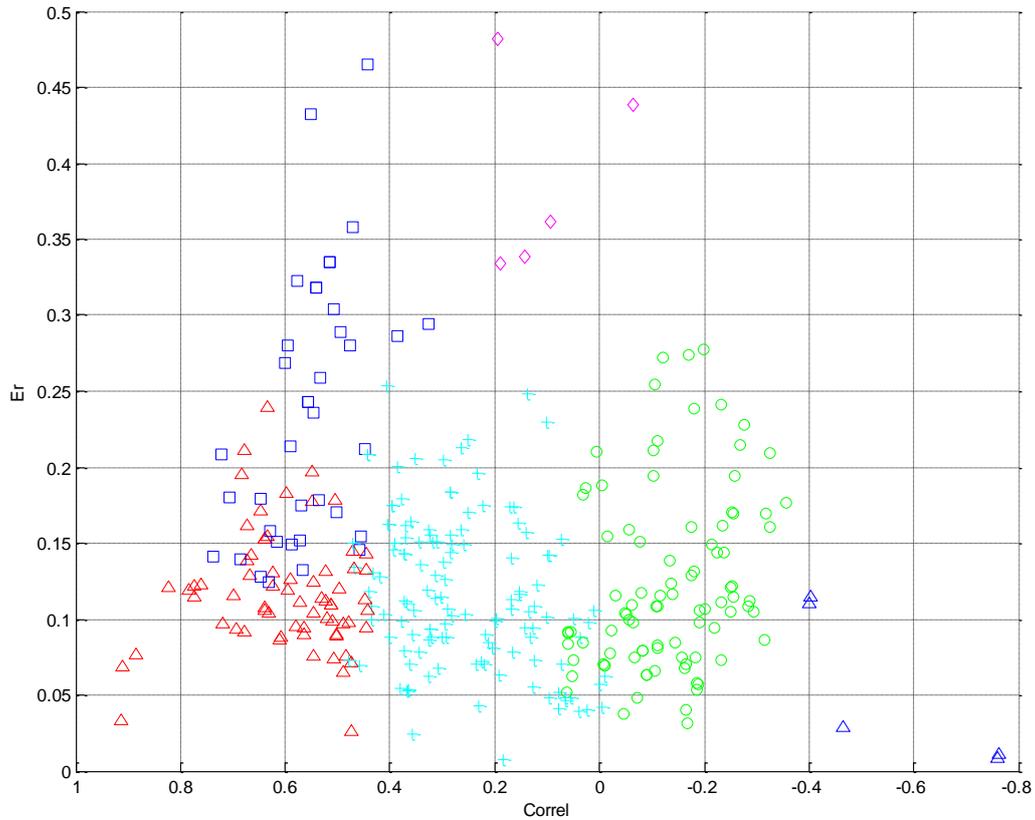


Figure 25, Clustered Funds – Correlation vs Expected Return

Unfortunately it is not possible to display all return characteristics used in the clustering procedure in a scatter plot, so we limit ourselves to the 2D plot⁵⁷ and draw the two characteristics were the two small groups are separated best. With the line plot below it is possible to compare the clusters among all characteristics (one line stands for one fund):

⁵⁷ MATLAB provides also a very nice 3D plotting device, but without the possibility to rotate the picture it is hard to see things on a 2D surface.

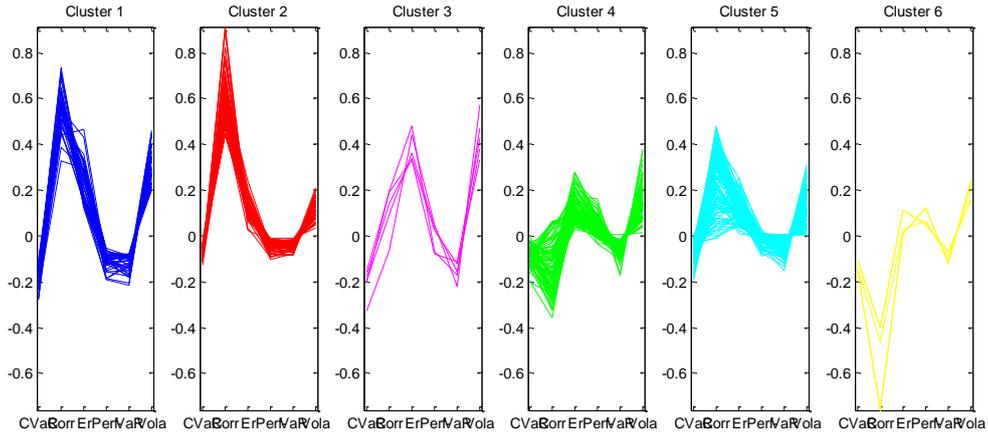


Figure 26, Clustered Funds – Line Plots

The most popular partitional clustering method is the **k-means** algorithm. The idea behind this technique is to assign each fund to a center and to choose the centers in such a way that the intra cluster variance is minimized. The following figure was produced with the kmeans command using the cityblock metric. Besides the distance metric one has to define the initial centers. As the k-mean algorithm is an iterative procedure it is sensitive on the choice of these initial centers. We used randomly drawn centers from the sample observations.

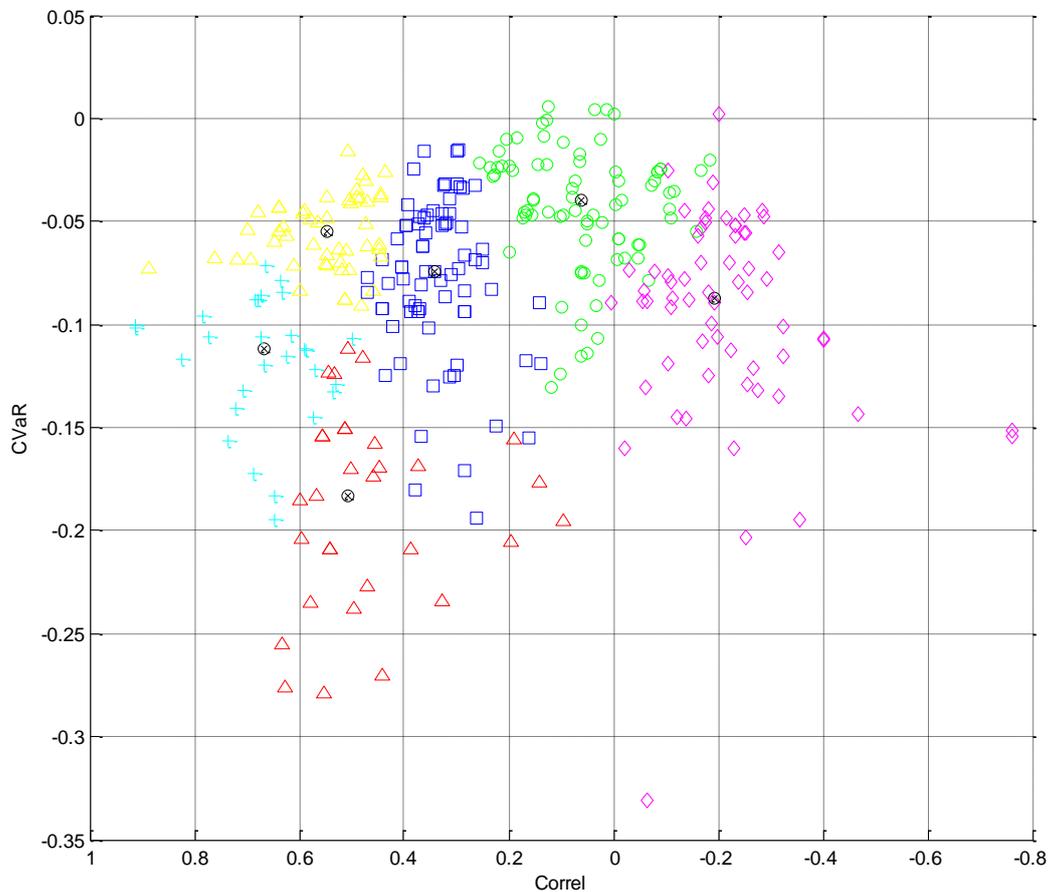


Figure 27, Clustered Funds – Correlation vs CVaR

The black encircled crosses are the centers. The fact that there is overlapping of the groups can be explained by closer similarity in other than the displayed characteristics. The advantage with the k-means algorithm is that it produces more equally big clusters and it is quite powerful when analyzing big data sets. As we have a limited amount of data and because of the intuitiveness of the dendrogram we recommend the hierarchical agglomerative clustering for our purposes. The clusters then can be used for peer group analysis for example.

There are also more modern ways of clustering that do not assign a fund to one specific cluster, so called fuzzy clustering. As the clusters found by the techniques above are not marked off sharply such techniques may make sense in the case of hedge funds and it would be an interesting research project to find out the proper specifications.

2.6.4 Individual Peer Indices

The initial notion was to come up with a technique to fill up a fund's track record based upon peer funds and not upon risk factors. We propose the following procedure:

- a) Find the n most similar funds that have a track record long enough.
- b) Create an index out of these funds.

Ad a: The first step will be to filter funds with the required minimum length. Then we propose not to set a fixed n , but rather to use all funds that satisfy certain requirements. This can be done in the following way (fund X is the fund whose track record you want to fill up):

- Calculate the proximity matrix for the qualitative data as described in section 2.6.2.
- Get the upper 10% percentile of similarity over all funds.
- Remember the funds that fall into this upper 10% in the column of fund X.
- Do the first three steps also for the return characteristics as described in section 2.6.3.
- Take the funds that are in the upper 10% of the qualitative *and* the quantitative criteria.
- Check if the beta of the selected funds to fund X are reasonable close to one.

Ad b: For the calculation of an index out of the selected funds we suggest to get the best one dimensional summary out of the data via the first principal component of a PCA like discussed in section 2.5.6.

3 Optimization

In this chapter first a literature overview of different risk measures considered for hedge fund data is provided. We critically discuss the different methods and recommend and motivate the usage of a certain combination of measures at this end of the section. In the next section we present different ways how to use linear constraints and propose ways to model more complex constraints. After this a flow diagram is provided to illustrate the complexity of integrating different filtering, data enhancement and optimization techniques. Then we provide the outcome of some first portfolio optimization experiments and conclude them by recommendations about certain input parameters and propositions for future research.

The dataset we get out of the data enhancement procedure next has to be optimized over the weights of the portfolio. A critical question is how to model the preferences of the person we do the optimization for. There are a lot of risk measures that try to capture these preferences and they will be discussed in the next section, but first we want to talk about how to quantify reward.

In general the **reward** will be measured as the expected return. This is the most natural and most frequently used measure and we will use it in this thesis, too. In the hedge fund literature only (Morton, Popova, & Popova, 2005) choose another approach, namely the probability to outperform a certain benchmark. Their paper will be discussed in section 3.1.5. It is also possible to take the reward not into account at all and just compute a minimum risk portfolio. This is considered by (Amenc & Martellini, 2002) and (Alexander & Dimitriu, 2004). See section 3.1.2 for details. The most correct way would be to derive the measure of reward from the incentive structure of the manager's contract and the reactions of his recent and possible future investors on certain events, reactions like a higher possibility of withdrawal if the fund is under water for a longer time.

3.1 Risk Measures

Unlike the reward risk is not so easy to define and even harder to measure. There is a whole body of literature on this theme and not all of the proposed measures can be discussed here. The focus will lie on popular risk measures and such that are frequently used to analyze hedge fund data.

3.1.1 Standard Deviation and Mean Variance Analysis

The best known risk measure is definitely the standard deviation or volatility (and equivalent to it in terms of portfolio allocation the variance). Together with the expected return as a measure of reward this leads to the so called mean variance optimization or μ, σ principle. If no constraints are imposed the portfolio optimization problem can be solved analytically, otherwise it can be solved by quadratic programming.

But there is a problem with the standard deviation or variance that can be described as follows: Markowitz justified his mean variance analysis together with Levy (Levy & Markowitz, 1979) as an approximation of a general utility function. This can be shown via the Taylor series expansion of the expectation of the **general utility function**:

$$\begin{aligned}
 E[U(R_p)] \approx & U(\bar{R}_p) + \frac{1}{2} \frac{\partial^2 U(\bar{R}_p)}{\partial R^2} \underbrace{E[(R - \bar{R}_p)^2]}_{\text{variance}} \\
 & + \frac{1}{3!} \frac{\partial^3 U(\bar{R}_p)}{\partial R^3} \underbrace{E[(R - \bar{R}_p)^3]}_{\text{skewness}} + \frac{1}{4!} \frac{\partial^4 U(\bar{R}_p)}{\partial R^4} \underbrace{E[(R - \bar{R}_p)^4]}_{\text{kurtosis}} \\
 & + \sum_{n=5}^{\infty} \frac{1}{n!} \frac{\partial^n U(\bar{R}_p)}{\partial R^n} \underbrace{E[(R - \bar{R}_p)^n]}_{\text{higher order moments}}
 \end{aligned}$$

Figure 28, (Lhabitant, 2004, S. 279), Formula Taylor Series Expansion of a General Utility Function

where $U(\cdot)$ is the utility function, R the returns and \bar{R}_p the mean portfolio return. If variance would be an optimal criterion either all deviations higher than two must be zero (this implies a quadratic utility function) or all moments higher than two must be zero (this implies normally distributed returns). As shown in 2.2 the returns of most of the funds in our databases are not normally distributed and from our point of view a quadratic utility function is not appropriate to reflect the utility of a typical investor. Nonetheless mean variance optimization is so frequently used that we have to calculate its outcomes at least as a benchmark for the other techniques.⁵⁸

As the mean variance optimization is the most commonly used technique for portfolio optimization it is also the most heavily researched. As the quality of every optimization algorithms output depends mainly on its input there was put a lot of effort in **estimating** the **parameter** needed in the mean variance optimization. To discuss all the techniques would go beyond the scope of this thesis therefore hedge fund related papers that use certain methods will be discussed only very briefly:

⁵⁸ Details about the mean variance optimization can be found in every finance textbook.

As there seems to be predictability in lagged risk factors, one can use it to forecast the expected return, or at least to get an idea of the future direction. (Amenc, Bied, & Martellini, 2003) do so and find that there is much value in using these forecasts in asset allocation decisions. A prominent way to come up with a guess about the expected return is the model of (Black & Litterman, 1992). (Martellini, Vaissie, & Ziemann, 2005) extend this approach to higher moments and use it in the case of hedge funds. They also consider vector autoregressive models and report considerable improvements. (Giamouridis & Vrontos, 2007) focus on the use of different estimation approaches for the variance. They conclude that the use of dynamic models improves the out of sample risk adjusted return.

Besides the fact that mean variance optimization is only correct with normal distributed returns or quadratic utility there is also other critique, namely that it works as an **error maximizer**, as remarked in (Michaud, 1989) and (Britten-Jones, 1999). And indeed the outcome of the mean variance optimizations depends strongly on its input – the expected returns and the covariance matrix. If there is an estimation error that leads to an extraordinary high expected return or diversification benefit the algorithm will pick this stock and weigh it heavily. As errors come and go over time the mean variance optimization produces erratic changes in portfolio weights, not executable in stock markets and least of all in the hedge fund world with its long lock ups and redemption periods. The effect of erratically fluctuating weights can be observed not only with mean variance optimization but also with most of the other optimization techniques, too. Nonetheless most of the methods to get more robust results were developed for the mean variance case. Therefore we will discuss them in the next section before dealing with other optimization techniques.

3.1.2 Sources of Robustness

One suggestion to overcome the problem of erratic outcomes is to calculate just a minimum risk portfolio. With the variance as risk measure this will be the well known minimum variance portfolio **MVP**. The big advantage of this approach is that one no longer has to care about the uncertainty in the expected returns but just deal with covariances, which can be estimated more reliably. The drawback is that in a realistic problem setup always a certain level of return will be specified.

(Amenc & Martellini, 2002) perform such an analysis with hedge fund indices. They find that their MVP had similar out of sample expected return but much lower volatility than the equally and the value weighted portfolio. Relative Value and arbitrage strategies were over emphasized, emerging

markets, global macro and long/short equity not included. (Alexander & Dimitriu, 2004) also compute the MVP but do not compare it to a mean variance portfolio. They focus more on the impact of covariance shrinkage techniques. If one does not want to abstain from the expected returns it is possible to shrink them towards an arbitrary “grand mean” or towards the return of the MVP via Bayes-Stein Estimation⁵⁹.

Normally the estimation of the covariance matrix is assumed to be more reliable than the one for the expected returns, but for a large number of funds and short time series severe problems occur. For example if one wants to estimate a covariance matrix for 2000 funds with a two year estimation window one has only 48,000 data points to estimate nearly two million parameters.

To get more reliable estimators for the covariance matrix there exist also several **shrinking** methods for the covariance matrix.⁶⁰ The tradeoff with these shrinkage techniques is always between specification and estimation error. (Amenc & Martellini, 2002) perform a principal component analysis and shrink the covariance matrix to the factor matrix of the significant principal components. (Alexander & Dimitriu, 2004) perform a very similar study and find that with a non negativity constraints in place the usage of the sample covariance matrix led to a better out of sample performance than the usage of the cleaned covariance matrix.

These findings are in line with the research of (Jagannathan & Ma, 2003). They find that to **imply constraints** in the optimization, the method that is most frequently used by practitioners, produces very similar outcomes (in the sense of robustness) to shrinking the covariance matrix. Normally maximum weights for a single asset are defined and maximum and minimum weights for the different asset classes. How to set meaningful constraints in the hedge fund world will be discussed in section 3.2.

Another method oriented on application is to employ a larger number of different data enhancement and optimization techniques. Then the outcomes can be averaged. Or, to get another source of robustness, one can first calculate the rank of the portfolio weights suggested by the different techniques and then average over the ranks. This approach helps to avoid misleading results in one technique.

⁵⁹ For details see (Jorion, 1986) and the initial (James & Stein, 1961).

⁶⁰ Methodological very good papers on the subject are (Ledoit & Wolf, 2003), (Ledoit & Wolf, 2004) and (Chan, Karceski, & Lakonishok, 1999).

The final method to get more robust outcomes with mean variance optimization (or with portfolio optimization in general) we want to introduce is the **resampling** technique as discussed in (Michaud, 1989) and (Scherer, 2002). One possible procedure works as follows: draw random outcomes for the asset returns (according to their correlation structure using the Cholesky decomposition) hold everything else constant (especially the covariance matrix) and optimize. Then repeat this procedure several times and calculate the mean of the resulting portfolio weights. A variation is that also the covariance matrix can be drawn randomly, for example from different estimations. The reason why this relatively simple technique is considered only relative recently is because it affords a lot of calculation power.

(Haberfelner & Kappel, 2007a) performed a study using a variation of the resampling concept. The aim was to optimize hedge fund styles to come up with a recommendation for the tactical asset allocation of a fund of hedge funds. The problem when simply optimizing hedge funds is (besides the erratic fluctuations in the weights) that the hedge funds **indices** include hundreds of funds. A normal fund of hedge funds on the other hand can only include at most 5 to 10 funds to represent a certain strategy. But this means that the indices, because of better diversification, have a much better risk return profile than the 5 funds that should represent the strategy in the fund of hedge funds portfolio. Calculating weights with inputs that then cannot be mimicked clearly does not make sense.

The idea to overcome this problem was to draw randomly 5 funds for each strategy and creating an index out of them, optimize these created indices and repeating the procedure 40 times⁶¹. In a sense this was resampling the index returns and covariance matrix simultaneously. The following figures show the outcome:

⁶¹ It turned out that the tradeoff between the standard error in the averaged weights and computational effort was best with this number of iterations.

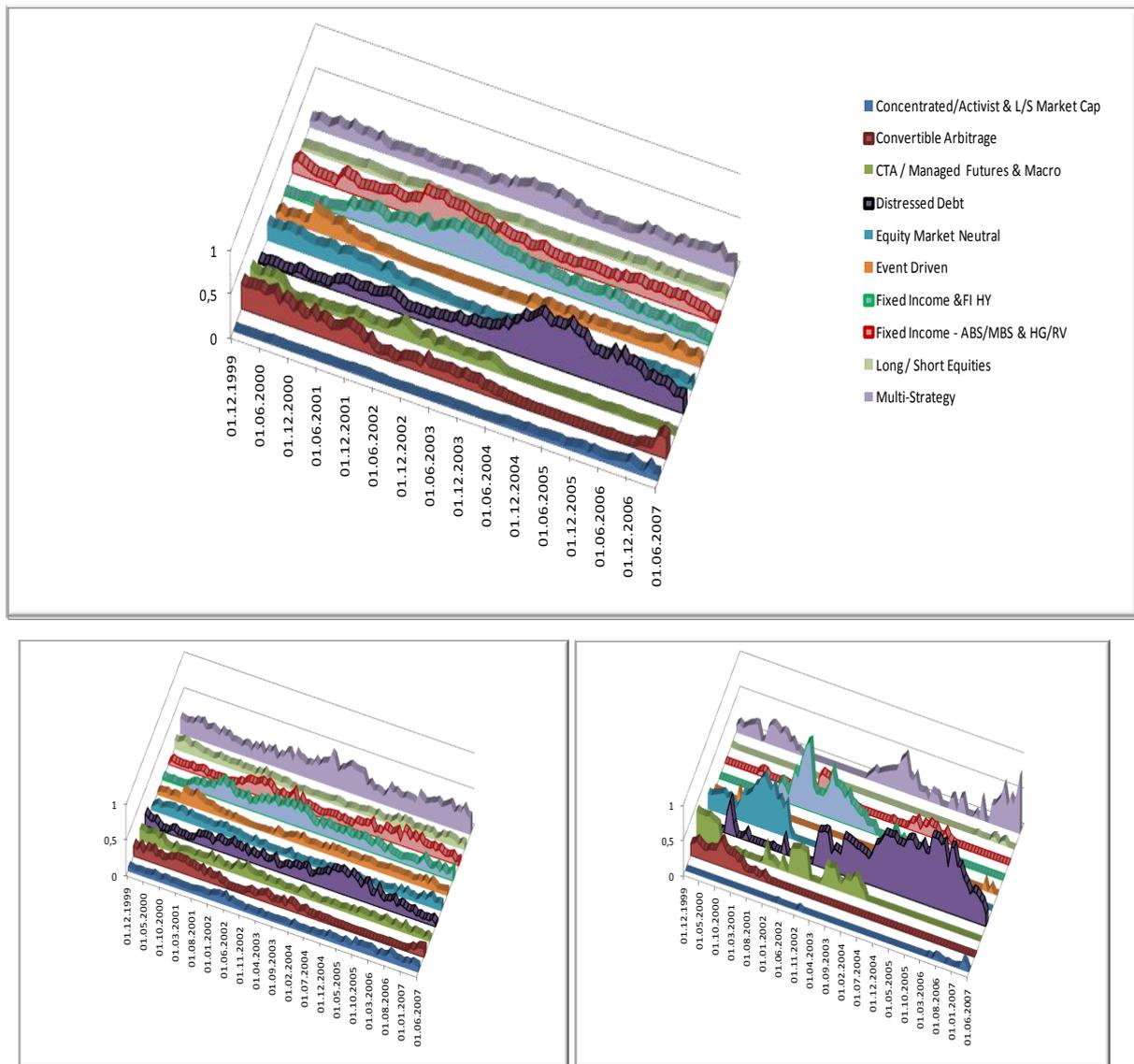


Figure 29, Resampling – Hedge Fund Strategie Weights from 1999 to mid 2007

The first figure shows the weights when calculated according to the procedure described above. The second figure shows the outcome when optimized on fund level (without first calculating an index and then optimizing). The third one was calculated without using the resampling technique, calculating a $1/n$ index for each strategy.

The outcomes produced with the resampling technique are much smoother and much better dedicated to be used for real investment decisions. An interesting fact is also that the multi strategy funds get a much higher weight in the “resampling + optimization on fund level” (second figure) approach than with the “resampling + calculating an index + optimization on index level” approach. This may be because they are so diversified among themselves. It is also nice to see that the resampling technique was able to capture the “return of convertible arbitrage”, whereas on the index level the optimizer was not able to catch this phenomenon.

3.1.3 Value at Risk

The value at risk (VaR) summarizes, loosely speaking, “the worst loss over a target horizon that will not be exceeded with a given level of confidence” (Jorion, 2007). With the frequently used delta normal method the assumptions are normally distributed risk factors and linear dependencies. As argued in section 2.1 these assumptions do not hold true for hedge funds. Besides this it is not clear at all what the risk factors for hedge funds really are.⁶²

When calculating the VaR directly from the fund’s return distribution as $z \cdot \sigma$, where σ is the fund’s annualized standard deviation and z confidence level, the assumption is normally distributed fund returns. One method to account for non normality is the so called (Cornish & Fisher, 1937) expansion. Instead of the confidence level from the normal distribution, z , the volatility gets multiplied by

$$z_{CF} = z + \frac{1}{6}(z^2 - 1)S + \frac{1}{24}(z^3 - 3z)K - \frac{1}{36}(2z^3 - 5z)S^2$$

where S is the skewness and K the (excess) kurtosis of the distribution. Another way can be to calculate the VaR from the discrete historical distribution.

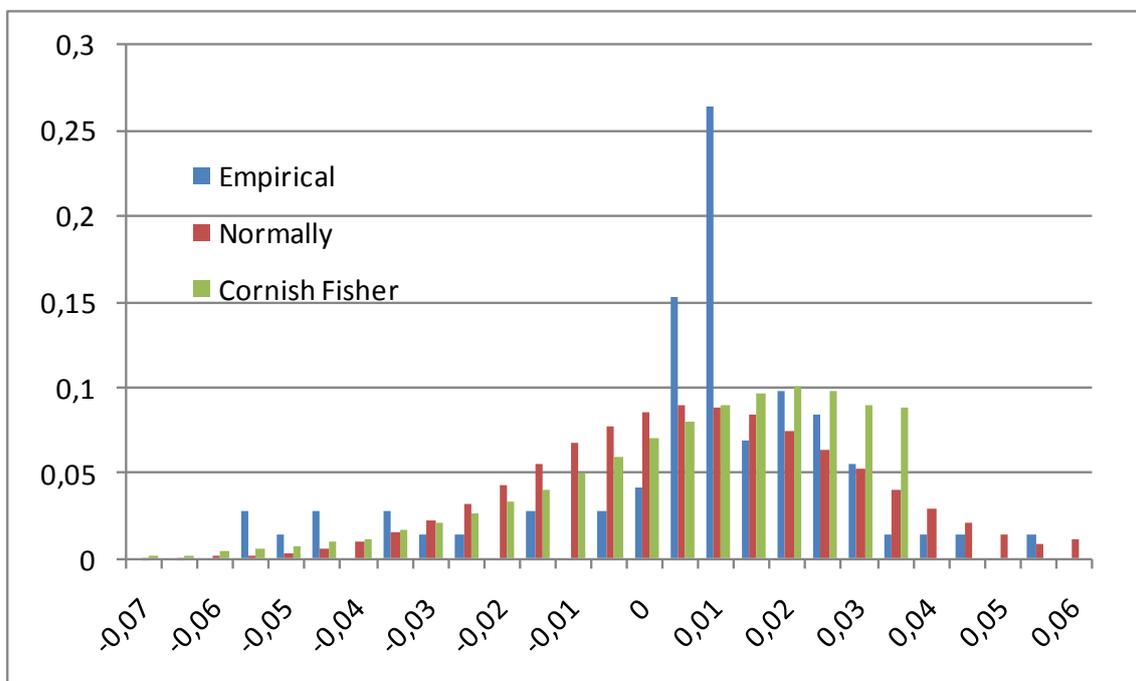


Figure 30, VaR Measures – Empirical, Normally Distributed and Cornish Fisher

⁶² If we write back track records with the factor model approach described in 2.5.7 we do a modification of the delta - historical simulation method. When accounting for nonlinearity, this essentially tries to capture the same problem like the delta-gamma method.

The according VaR numbers are:

	VaR 90%	VaR 95%	VaR 99%
empirical	-0,036	-0,052	-0,058
standard	-0,028	-0,037	-0,052
CF	-0,030	-0,043	-0,068

The Cornish-Fisher (CF) VaR always exceeds the VaR computed under the assumption of normally distributed returns by construction if the distribution has negative skewness and positive kurtosis. Nonetheless it still underestimates the risk in this (quite extreme, but real) example. Only for the 99% VaR the Cornish-Fisher VaR is lower, but with 72 observations the 99% empirical VaR is just the smallest return for the empirical distribution. The CF expansion seems to be made for the estimation of left tail risk, the right tail looks a little bit weird.

But there are three problems with the VaR that are the same for all three approaches:⁶³

- It lacks of **subadditivity**, one of the properties required for a coherent risk measure in the sense of (Artzner, Delbaen, Eber, & Heath, 1999).⁶⁴ This can be seen from a simple example: “Consider two independent investment opportunities each returning a \$1 gain with probability 0.96 and \$2 loss with probability 0.04. Then, 0.95-VaR for both investments are 1. Now consider the sum of these two investment opportunities. Because of independence, this sum has the following loss distribution: -\$4 with probability $0.04 \cdot 0.04 = 0.0016$, -\$1 with probability $2 \cdot 0.96 \cdot 0.04 = 0.0768$, and \$2 with probability $0.96 \cdot 0.96 = 0.9216$. Therefore, the 0.95-VaR of the sum of the two investments is -1, which exceeds 2, the sum of the 0.95-VaR values for individual investments.” This means VaR may discourage diversification.
- VaR may provide an inadequate picture of the risk as it does not measure the losses that exceed the VaR. Therefore a minimization of the VaR may lead to stretch the **tail exceeding the VaR**.
- The VaR is **difficult to optimize**, it is a non-smooth, non-convex function of the portfolio weights that has multiple local extrema.

⁶³ Confer (Rockafellar & Uryasev, 2000)

⁶⁴ The other properties are monotonicity, positive homogeneity and translation invariance.

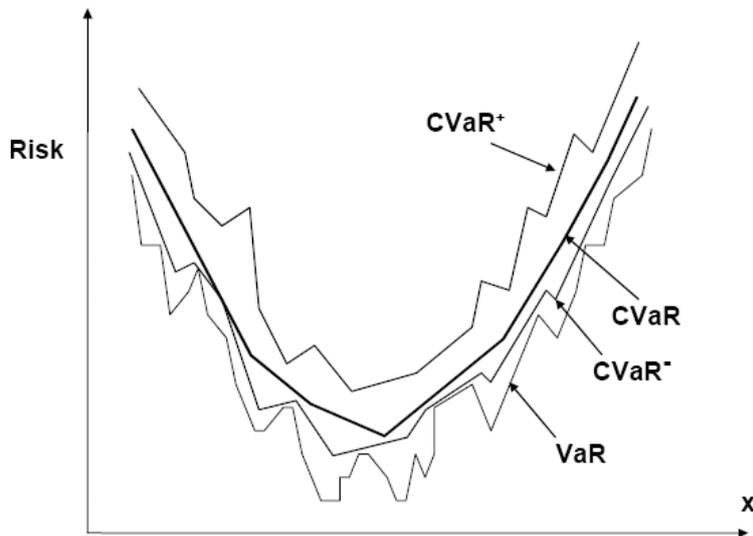


Figure 31, (Uryasev), Conditional Value at Risk – $CVaR^+$ and $CVaR^-$ are the expected losses strictly (weakly) exceeding the VaR. The CVaR is a weighted average of VaR and $CVaR^+$.

A measure that solves this problem is the Conditional Value at Risk (CVaR), also called Mean Excess Loss, Mean Shortfall, Expected Shortfall or Tail VaR.⁶⁵ The CVaR is defined as the expected loss given that the loss exceeds the VaR. Although it could be calculated from a parametric distribution, too,⁶⁶ in the multivariate case the CVaR gets calculated best from a set of scenarios with a simulation approach. The easiest way is to use the discrete historical distribution.

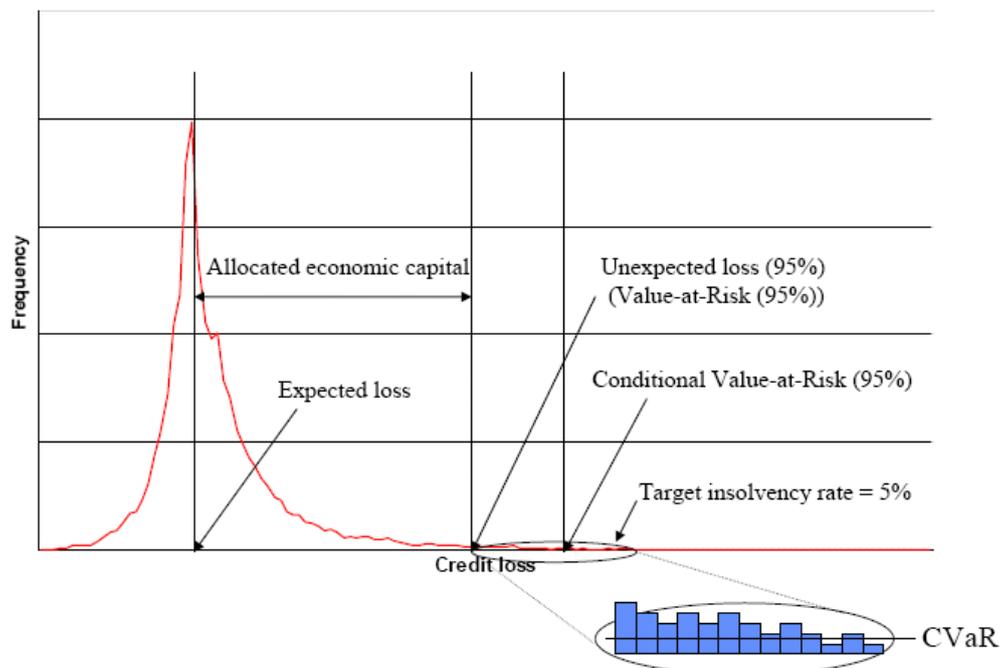


Figure 32, (Uryasev), Risk Measures
(Note that a loss function is displayed, therefore the negative outcomes are on the right tail)

⁶⁵ It was shown by (Pflug, 2000) that CVaR is a coherent risk measure.

⁶⁶ See (Andreev, Kanto, & Malo, 2005) for details.

Another very nice property of the CVaR is that it can be reduced to a set of linear constraints and therefore optimized very efficiently by large scale linear optimizers.⁶⁷

The CVaR is defined as

$$\text{CVaR}_\alpha(x) := \frac{1}{1-\alpha} \int_{f(x,y) \geq \text{VaR}_\alpha(x)} f(x,y)p(y)dy \quad (1)$$

Note that $f(x,y)$ is a loss function. Adding a constant before the integral and subtraction is after the integral in (1) does not change the outcome:

$$F_\alpha(x, \gamma) := \gamma + \frac{1}{1-\alpha} \int_{f(x,y) \geq \gamma} (f(x,y) - \gamma) p(y)dy \quad (2)$$

This can be rewritten as

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{1-\alpha} \int (f(x,y) - \gamma)^+ p(y)dy \quad (3)$$

“where $a^+ = \max\{a,0\}$. This function, viewed as a function of γ , has the following important properties that make it useful for the computation of VaR and CVaR:

1. $F_\alpha(x, \gamma)$ is a convex function of γ .
2. $\text{VaR}_\alpha(x)$ is a minimize over γ of $F_\alpha(x, \gamma)$.
3. The minimum value over γ of the function $F_\alpha(x, \gamma)$ is $\text{CVaR}_\alpha(x)$.

As a consequence of the listed properties, we immediately deduce that, in order to minimize $\text{CVaR}_\alpha(x)$ over x , we need to minimize the function $F_\alpha(x, \gamma)$ with respect to x and γ simultaneously” (Cornuejols & Tütüncü, 2006, S. 276-277).

For a discrete empirical probability distribution based on scenarios (3) can be rewritten as

$$\tilde{F}_\alpha(x, \gamma) := \gamma + \frac{1}{(1-\alpha)S} \sum_{s=1}^S (f(x, y_s) - \gamma)^+ \quad (4)$$

The according optimization problem is:

$$\min_{x \in X, \gamma} \gamma + \frac{1}{(1-\alpha)S} \sum_{s=1}^S (f(x, y_s) - \gamma)^+ \quad (5)$$

Figure 33, (Cornuejols & Tütüncü, 2006, S. 276-277), Formulae of the Linearization of the CVaR 1

⁶⁷ Confer the initial paper by (Rockafellar & Uryasev, 2000) or the didactically very good book by (Cornuejols & Tütüncü, 2006, S. 273-280).

The trick here is: If the γ gets smaller the weighted sum in (5) grows faster for a small α and grows slower with a high α . The left part (this is only the γ) is not dependent of α . So there is a tradeoff determined by the α between the γ and the weighted sum. The point where the increase (decrease) in γ is exactly offset by the decrease (increase) in the weighted sum, there is the minimum. In the minimum the γ is exactly the $VarR_\alpha$. If the γ is optimal and the x are not the overall outcome of (5) will not be optimal, a change in x will improve the outcome. If the x are optimal for a given non optimal γ a change in γ will improve the outcome (maybe the x have to be changed too).

By replacing the $(f(x, y_s) - \gamma)^+$ in (5) by an auxiliary variable z_s the problem is reduced to a set of linear inequalities:

$$\begin{aligned} \min_{x, z, \gamma} \quad & \gamma + \frac{1}{(1-\alpha)S} \sum_{s=1}^S z_s \\ \text{s.t.} \quad & z_s \geq 0, \quad s = 1, \dots, S, \\ & z_s \geq f(x, y_s) - \gamma, \quad s = 1, \dots, S, \\ & x \in X. \end{aligned} \tag{6}$$

Figure 34, (Cornuejols & Tütüncü, 2006, S. 278), Formulae of the Linearization of the CVaR 2

We therefore strongly recommend using the CVaR out of the family of VaR related measures. A MATLAB code for portfolio optimization with CVaR constraints can be found in appendix 5.6 and in appendix 5.7.

3.1.4 Higher Moments

As argued above the CVaR approach is based on scenarios and the parametric mean variance optimization is only applicable in a meaningful way under the assumption of normality. But there are other parameterized approaches that account for non-normality.

The approach of (Davies, Kat, & Lu, 2005) uses polynomial goal programming to simultaneously optimize the investor's preferences on variance, skewness and kurtosis. The problem with this approach is that it is not clear how to set the preference coefficients for the variance, the skewness and the kurtosis⁶⁸ and how to optimize the cubic function required to maximize the kurtosis.

Another approach is considered by (Jondeau & Rockinger, 2006). They expand a CARA (Constant Absolute Risk Aversion) utility function with the Taylor series and then calculate a closed form solution for the portfolio allocation problem. The drawback from our point of view is again that we

⁶⁸ They are defined as the difference from an optimal minimum variance, minimum skewness and minimum kurtosis portfolio.

neither know if a CARA function really reflects the utility of the investor adequately nor do we know her risk aversion coefficient, if CARA is sufficient. In addition to this Jondeau and Rockinger do not include constraints in the sense of maximal and minimal allocations in their problem set up.

A problem that goes along with all parameterized approaches is that every time higher moments are considered the aspect that should be captured is the tail risk. But the tail risk is better captured by the CVaR from our point of view. Another shortcoming is that higher co-moments are even more unstable as mean and covariance. Especially the odd moment, the co-skewness, is hard to capture.

3.1.5 Other Techniques

Achieve a benchmark

(Morton, Popova, & Popova, 2005) do not use the expected return as reward but choose another approach. They argue that the reward for the fund manager is not the (risk adjusted) return but “due to the manner in which today’s portfolio managers are judged, we regard the probability of achieving the benchmark [like the S&P500 or the HFR Fund of Funds] as the reward measure and we use expected regret [the magnitude by which the benchmark is missed] as the measure of risk.” The expected utility (a weighted sum of reward and regret) then can be optimized (after reformulation) as a mixed-integer linear program. Besides the fact that MATLAB does not provide a MILP solver, (Morton, Popova, & Popova, 2005) state themselves that neither their “utility function nor its expectations are concave functions. So, our utility function does not satisfy the notion of coherence” in the sense of (Artzner, Delbaen, Eber, & Heath, 1999).

Omega

A very interesting risk measure is the so called Omega (Keating & Shadwick, 2002). Essentially it is the probability weighted gains divided by the probability weighted losses. Gains and losses are both considered with respect to an arbitrary threshold.

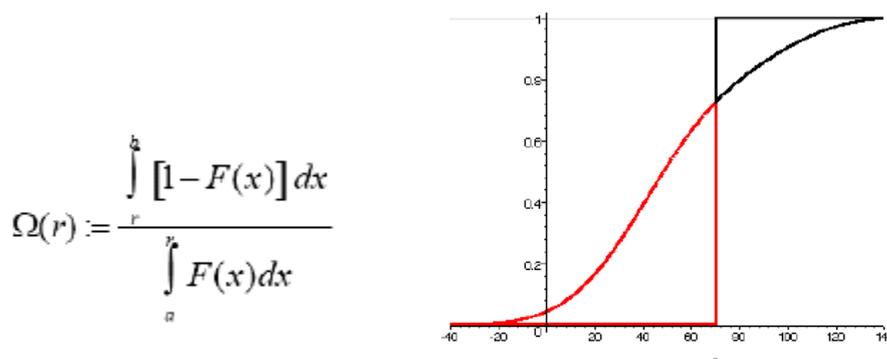


Figure 35, (Keating & Shadwick, 2002), The Omega Risk Measure

By shifting the threshold a graphic can be created, where several assets are very well comparable:

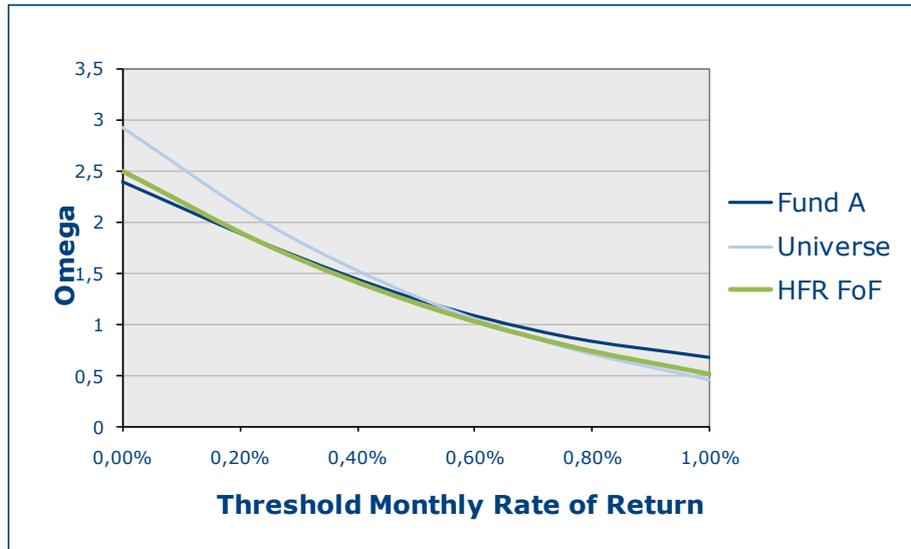


Figure 36, (Slightly adapted from an Excel Sheet used by Benchmark Capital Management), Omega vs Threshold

Mean Absolute Deviation

Instead of the quadratic distances (as with the variance) the absolute distances are considered with the mean absolute deviation (MAD). The MAD has the nice feature that it can be transformed into a set of linear inequalities (Konno & Yamazaki, 1991), very similar to the CVaR, and therefore can be optimized also for a large number of assets with linear optimizers. In the case of a multivariate normal distribution the outcomes of the mean variance and the mean MAD optimization are identical⁶⁹, and it can be shown that the mean-MAD portfolios are never stochastically dominated (Ruszczynski & Vanderbei, 2003). Therefore we recommend replacing the constraint on the variance in a “mean-variance optimization with CVaR constraint” with the MAD for applications with a lot of funds. The according MATLAB code can be found in appendix 5.7.

Maximum Drawdown

The maximum draw down is the maximum drop from a historical peak. More loosely speaking it is the maximum loss an unlucky investor can make (buying at the peak and selling at the valley). The special thing about maximum drawdown is that it is a loss measure with memory, as the last peak has to be kept in mind. A light variation is the conditional draw down, the average of the x% highest draw downs. It can be reduced to a set of linear in-(equalities) and is convex in the portfolio weights and therefore can be optimized in a similar manner as the CVaR (Krokhmal, Uryasev, & Zrazhevsky, 2002).

⁶⁹ The MAD can then be calculated as $\sqrt{\frac{2}{\pi}} \sigma$

3.1.6 Conclusion

Clearly the most appropriate way to optimize a fund of hedge funds portfolio would be to determine the client's exact preferences and then optimize with respect to this function. If the client is the fund of funds manager an approach similar to the one of (Morton, Popova, & Popova, 2005) could be used, although there would be some refinements to do. If the client is an average investor or a fund of fund manager who wants a tool that can be used for marketing reasons as well the best way would be to do mean-variance optimization with a constraint on the VaR, because when average investors talk about risk, they are used to express it in term of volatility and Value-at-Risk. If they have worked with these measures for a longer time they get a quite good feeling about what level of risk they can tolerate expressed in these measures. And in the end this setup is not the worst one as there is a measure of dispersion and one that controls for asymmetry.

The only challenge is to put it in a way that optimization algorithms can handle the problem. As argued above there are good reasons to replace the VaR by a CVaR constraint, reasons that can be communicated to investors, too. The other adaption would be to use the MAD instead of the volatility in cases with a large universe, as the MAD is linearisable and therefore much faster to optimize than the volatility. Another advantage is that the mean – MAD/CVaR optimization is based on scenarios (stochastic programming), which makes it very flexible in determining dependency structures.

3.2 Limit Architecture

Frequently there are constraints for the portfolio optimization coming for considerations of the fund of fund manager or form bounds defined in strategic asset allocation. But to imply constraints also makes sense for **robustness** reasons: it helps to avoid corner solutions and leads to a better tradeoff between specification and sampling error, very similar like the shrinkage of the covariance matrix (Jagannathan & Ma, 2003).

3.2.1 Linear Constraints

Normally group constraints will be imposed on risk blocks in the sense that only a certain proportion of the portfolio is allowed to be exposed to a certain risk factor. This risk block can be:

- Regional risks
- Market risks
- Hedge fund strategies
- Liquidity
- Leverage

The key determinant here clearly is a carefully acquired matrix that contains the risk exposures of the funds. This matrix is normally generated in a more qualitative process by reading fund prospects and talking to managers. Information about a funds “hedge fund strategy”, its regional exposure and its liquidity terms can be found in most databases.

$$\text{exposure vector}' * \text{weights} \leq \text{limit}$$

The exposure vector for “hedge fund strategy: convertible arbitrage” may contain zeros and ones depending on if the fund is convertible arbitrage or not. The limit is in terms of “percent of the portfolio”. If one want to set a limit on the *number* of funds this would require mixed integer linear programming (MILP). With other buckets like regional and market risks one may want to assign a fund not only to a single attribute, as a fund may act on different markets and in different regions. To account for this is no problem for the technical side; the exposure vector then gets filled with the according fraction, e.g. the entry for the “Europe” vector for one specific fund could be 0.4 and for the “US” vector 0.6. If one assumes leveraged exposures the entries must not even sum up to one. To define subordinate buckets (like “equity-like” and “bond-like” for the hedge funds strategies) is also no problem, this means simply to create two new exposure vectors. As linear programming is efficient also for very large optimization problems, one must not care about the number of exposure vectors.

As it is not possible to short hedge funds and as the usage of leverage for a fund of hedge funds is quite unusual a non-negativity and a full-unlevered-investment constraint are useful. To ensure that the portfolio does not get too concentrated and that a certain minimum number of funds is guaranteed a upper bound for the individual fund weights is advisable, too.

$$weight_i \geq 0 \quad \forall i$$

$$weight_i \leq upper\ limit \quad \forall i$$

$$\sum weights \leq 1$$

Another useful constraint is to set the weight of funds that are already in the portfolio to their actual values if they should stay in the portfolio. A similar constraint can be imposed for funds that are planned to enter the portfolio with a certain target weight because of more qualitative reasons. The optimization algorithm then is able to allocate new funds around a core portfolio.

$$weight_i = targetweight_i$$

Constraints on the turnover can be implemented in the following way:

$$\begin{aligned} \max(actualweight - transactionlimit, 0) &\leq weight_i \\ &\leq \min(actualweight + transactionlimit, upper\ limit) \end{aligned}$$

A constraint especially suitable for a fund of hedge fund is a market neutrality constraint. As the beta of a portfolio is just the sum of the betas of its constituents this is also a simple linear constraint.

$$-0.1 \leq \sum_i (beta_i * weight_i) \leq 0.1$$

A constraint normally used for mutual funds is a tracking error constraint. If a fund of hedge funds is designed to map a certain hedge fund index this makes sense in the hedge fund case, too. Unfortunately the tracking error is a quadratic function and therefore would require quadratic programming. Fortunately there is a way to replace it by a set of linear equations.⁷⁰

3.2.2 Logical and Cardinality Constraints

An interesting set of constraints are the ones that model logical conditions. There may occur the case that if one investment is made, another one has to be made, too. If we define x_1 and x_2 as integer variables, being one if the investment is made and zero if not we can model the problem as

$$x_1 - x_2 \leq 0.$$

⁷⁰ For details see (Cornuejols & Tütüncü, 2006, S. 181-184).

The restriction “if one investment is made the other one cannot” would look as follows

$$x_1 + x_2 \leq 1.$$

Restrictions like “in the portfolio 50% of the funds are allowed to have a lockup of over one year, but only if 80% out of these funds have a volatility smaller than 0.1” can be handled by defining three binary vectors, one for “has lockup above one year”, one for “has volatility smaller than 0.1” and one for “is in the portfolio”. The three vectors have to be multiplied element wise and the elements of the resulting vector have to be summed up. This sum has then to be divided by the sum of elements of the “is in the portfolio” vector. If this number is smaller than 0.8*0.5 the constraint is fulfilled. The only aspect that would require integer variables that change with the optimization process is the “is in the portfolio” vector. If one replaces “50% of the funds” and “80% out of these funds” by “50% of the portfolio” and “80% of this fraction” the problem is a simple linear constraint.

$$\sum_i (vola_i .* lockup_i .* weight_i) \leq limit_{lockup} * limit_{vola}$$

were .* is the element wise multiplication, $vola_i$ the indicator if the fund was volatility above 0.1 or not, $lockup_i$ the indicator for lockup above one year and $weight_i$ the weight of the fund in the portfolio (or the indicator whether the fund is in the portfolio or not).

A minimum number of funds in the portfolio can easily be ensured by setting an upper bound of the fund weights. To ensure that a maximum number of funds is not exceeded is not so easy. Just setting a lower bound will not work if the number of funds exceeds $1/lower\ bound$. Instead a limit has to be defined in the form of “only if the fund is in the portfolio, it must have a minimum weight”. A similar constraint is the minimum transaction lot: only execute a change in portfolio weights if the change is higher than for instance 1%. In the first case an indicator is required whether a fund is in the portfolio or not, in the second one if the size of the transaction is higher than 1% or not. These problems can be solved by a branch and bound algorithm, that creates a tree by setting a portfolio weight violating the conditional constraint either to 0 or to the bound (minimum weight and minimum transaction lot, respectively) at each node (Cornuejols & Tütüncü, 2006, S. 224-226).

Even harder the problem gets if one requires a fixed number of funds in the portfolio. This leads to the so called cardinality constrained portfolio optimization. To get a really exact solution all possible combinations of funds should be considered. The efficient frontier is then a patchwork of the

efficient frontiers of all the possible subsamples. For the case of two out of four assets we have $\binom{4}{2} = 6$ possible portfolios.

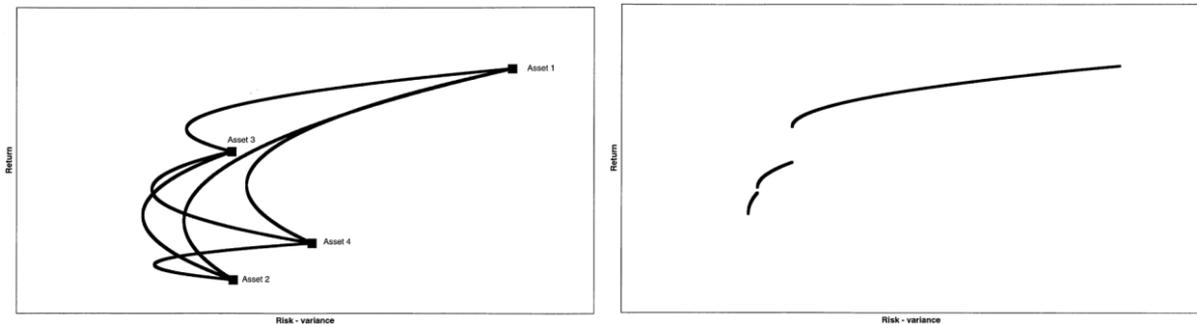


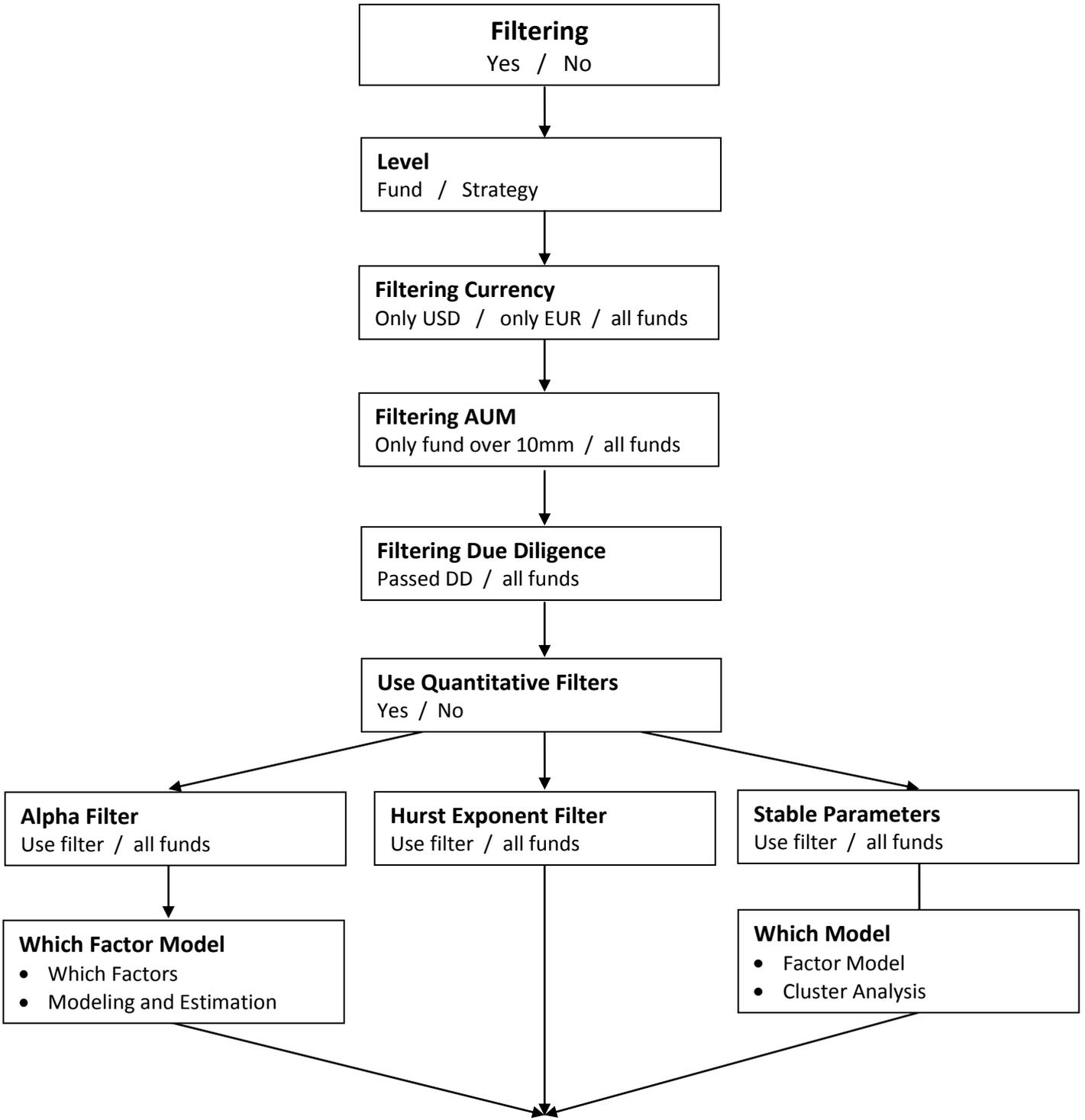
Figure 37, (Chang, Meade, Beasley, & Sharaiha, 2000), Cardinality Constrained Portfolio Optimization

For the realistic case of 30 out of 500 funds the number of possible combinations cannot even be calculated. Hence heuristics have to be used to get an approximately solution. (Chang, Meade, Beasley, & Sharaiha, 2000) test a genetic algorithm, a tabu search and a simulated annealing approach. They conclude that the best way is to use all three heuristics and to pool their results. The best performing heuristic was the genetic algorithm, with the drawback that it took the longest time to come up with a solution. It is also possible to use a branch and bound algorithm here and if the unbound solution contains a similar number of funds like the required number of funds the branch and bound algorithm is likely to perform much faster than the above mentioned heuristics.

To include such logical and cardinality constraints is clearly an interesting field for future research. Especially for a fund of hedge fund a minimum investment makes sense, as there are very high research and monitoring costs associated with each new fund.

3.3 Integration of Data Enhancement and Optimization Techniques

We now have introduced a large number of different filtering, data enhancement and optimization techniques. To get an idea of the complexity of the problem a little flow diagram is provided:



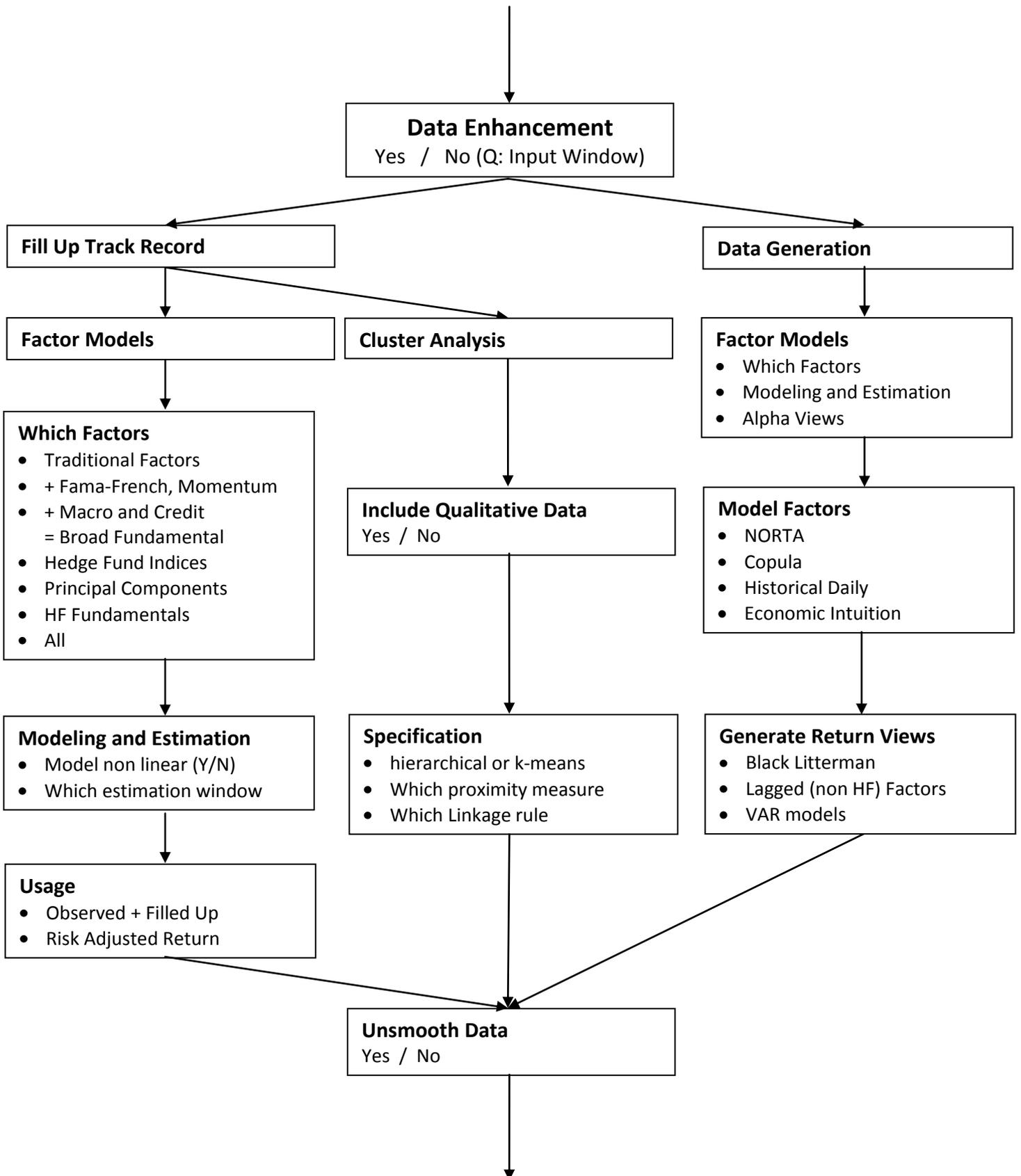




Figure 38, Flow Diagram

3.4 Portfolio Optimization Experiments

As the above flow diagram indicates it is simply impossible to assess all possible combinations of techniques and specifications, not only because of the high number of possible combinations but also because some of the techniques are quite time intense, especially simulation approaches. In the following section we provide the outcome of some first experiments that could identify at least directions of future research.

3.4.1 General Considerations

At the end all attempts will have to be assessed by the out of sample performance. But there arise several questions how to assess the out of sample performance:

The first one is to assume monthly, quarterly or annual **rebalancing**. Most realistic would be quarterly rebalancing. Also from the angle of portfolio turnover quarterly rebalancing may be advisable, as the fund of fund manager has more time to execute the scheduled adjustments. Nonetheless we will use monthly rebalancing in our experiments to get a better idea what is going on when using a certain technique.

The second topic is when to **apply** the **weights** from the optimization: The final performance figures for the hedge funds and the hedge fund indices come in normally around the 20th of the following month (i.e. the figures for February are available not until March 20th). So, even if the calculation, the decision and the execution process could take place in a few days (what is quite unrealistic) one would have to accept at least a one month lag. After discussions with practitioners we find that a three month lag is most realistic.

The third question that arises is what to do if a fund that was included in the optimized portfolio (optimization window $t-w$ to t) is no longer in the database in $t+1$. Because we have no information whether a fund stopped reporting voluntarily or if it was **liquidated** we propose to simply set the performance to zero and do not include the fund in the $t+1$ optimization. We argue this by assuming that the negative effects of having chosen a very bad fund that got liquidated and the positive effects of the investing in a very good fund that stopped reporting because it does not have to attract new investors offset each other in this specific month and that the money invested in this fund is freely available in the next month.

The last issue to consider are lock up periods and other liquidity terms of a hedge fund: The investor has to hold the hedge fund at least for the lock up period and after that lock up time she has to announce redemption some months in advance. As this restrictions influence the out of sample performance massively, they should be considered in a “real world” evaluation, too.

To have an idea about the value added (compared to throwing darts) of a certain method it is possible to compute different **benchmark** portfolios. The most simple is the 1/n portfolio with all funds alive. Clearly this is a very unrealistic portfolio as it holds at the end more than 2000 funds and is, because of the sheer number of included funds, highly diversified.⁷¹ To get more realistic benchmarks an approach of (Davies, Kat, & Lu, 2005) can be used: Randomly draw 5000 times 20 funds out of the sample and calculate certain risk parameters (they use mean, standard deviation, skewness and kurtosis) over a given period and then take the average of this parameters. Find the one portfolio out of the 5000 that best matches these parameters – this is our benchmark. Another reasonable approach is to use a fund of hedge fund index, as this benchmark will be free from biases for sure (Fung & Hsieh, 2002).

3.4.2 Quantile vs Dispersion Measures

The main reason why to use a quantile measure like VaR or CVaR as a constraint in an optimization is something that could be called tail drop: without controlling for the left tail the mean variance (or mean MAD) optimizer will pick the more left skewed asset out of two assets that have the same volatility because normally the tail risk gets compensated and the left skewed asset will therefore have a higher expected return.

To test if this is true we made the following calculations: Maximize the expected return for a given level of CVaR. Calculate the variance of the resulting portfolio. Maximize the expected return given the variance of the CVaR portfolio. Compare the CVaR and the expected return (the variance is the same). If the CVaR is worse⁷² and the expected return is higher with the mean-variance portfolio, the hypothesis is proven to be true. To have an indicator whether the distribution has a heavy left tail we

⁷¹ For the problem of over diversified benchmarks see 3.1.2.

⁷² As VaR measures are by convention calculated from loss distributions it should say higher here. But as the CVaR is calculated in this thesis from the ordinary return distribution it should say lower. To not fortify confusion we simple say worse.

also report the closed form CVaR for the normal distribution.⁷³ The following numbers were calculated with the HFR hedge fund indices:

MAX / MIN										
CONSTRAINT										
mean	0,0962	0,0988	0,1254	0,1331	0,1608	0,1643	0,1952	0,1957	0,2276	0,2284
stdev	0,041	0,041	0,0803	0,0803	0,1206	0,1206	0,1672	0,1672	0,2223	0,2223
CVaR	-0,01	-0,013	-0,025	-0,0339	-0,05	-0,0547	-0,075	-0,0775	-0,1	-0,1072
CVaR ~N	-0,0130879	-0,0129228	-0,0238279	-0,0233103	-0,0351837	-0,0349451	-0,049254	-0,04922	-0,0668415	-0,0667875

Figure 39, Table – Tail Drop Index Level

For each scenario the expected return is higher and the left tail is heavier with the mean – variance optimization. To get an idea how the resulting distributions look like, we provide the histograms for the distribution under CVaR constraint (-0.025) and the histogram for the according distribution under the volatility constraint (0.0803).

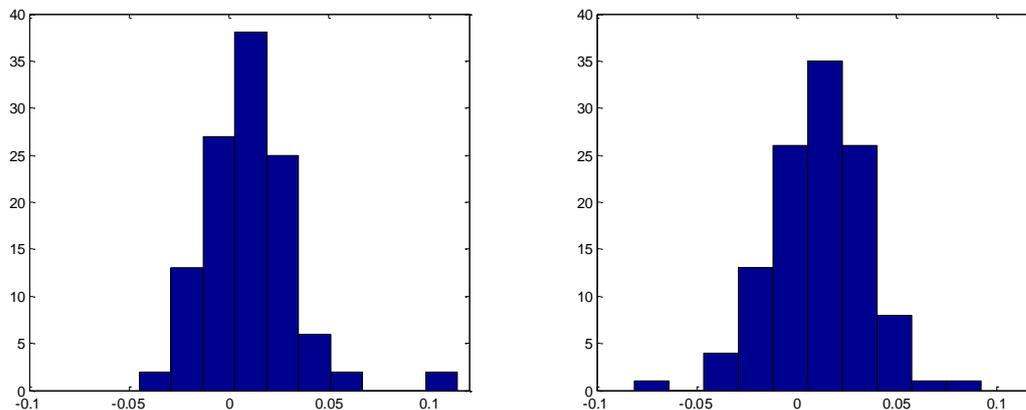


Figure 40, Distribution under CVaR Constraint and under Volatility Constraint

On fund level (merged HFR and Benchmark database, only funds that are denominated in USD and that have observations over the whole period) we use the MAD instead the variance, because of computational reasons. The outcome is similar:

MAX / MIN										
CONSTRAINT										
mean	0,2477	0,2642	0,2991	0,3099	0,3372	0,3469	0,3878	0,4036	0,4636	0,4673
stdev	0,0516	0,0564	0,0882	0,0908	0,1189	0,1273	0,1867	0,1951	0,3014	0,3002
MAD	0,0115	0,0115	0,0197	0,0197	0,0281	0,0281	0,0427	0,0427	0,0654	0,0654
CVaR	0	-0,0775	-0,01	-0,0233	-0,025	-0,0395	-0,05	-0,0547	-0,1	-0,1104
CVaR ~N	-0,0144527	-0,0149731	-0,0226384	-0,0228517	-0,0301656	-0,0323185	-0,0496211	-0,0515716	-0,0850226	-0,0843928

Figure 41, Table – Tail Drop Fund Level

⁷³ For the analytical calculation of CVaR under the assumption of other distributions see (Andreev, Kanto, & Malo, 2005).

Although the differences in the return are not too big there is a consistent pattern in all the scenarios. We can conclude that these numbers support the hypothesis of a tail drop. An interesting research question is if this phenomenon can be observed out of sample, too.

3.4.3 “Unconstrained” Optimization

Next we want to evaluate the influence of the time window used for the optimization (12, 24, and 36 months), the influence of the risk measure used and if it makes a big difference whether the optimized weights get applied one, two or three months later. Therefore we did several runs of a one month rolling unconstrained optimization. The term unconstrained is not totally exact as we applied a non negativity and a total investment/no leverage constraint. We did not use any data enhancement technique here; we just took all funds that have observations for the whole time window available with their original data. This means we started the first optimization at $b+36$, where b is June 1998, the beginning of our records. In one setup we optimized the data b to $b+36$, in the second one $b+13$ to $b+36$, in the third $b+25$ to $b+36$. In each setup we optimized with regard to all risk measures considered. Then we applied the weight calculated in the different setups one, two and three months later and stored the return. Then we shifted everything one month into the future and did the whole procedure again. At time t we had the optimization windows $t-35$ to t , $t-23$ to t and $t-11$ to t and applied the weights in $t+1$, $t+2$ and $t+3$. The risk measures we used were variance, MAD and CVaR. With the optimization we minimized the risk for a given target return to get comparable outcomes. We used three different target returns: 10%, 15% and 20%. For the calculation on fund level we just used MAD and CVaR, as the quadratic programming takes far too long time with more than 200 funds. With three risk measures, three time windows, three weight lags and three target returns we had 81 different portfolios. The numbers in the following tables are averaged over all parameters (time window, weight lag, risk measure and target return) but the one that is analyzed in the table.

As anticipated the out of sample performance is best when the weights get applied one month later. But interestingly these effects get offset by a decrease in the risk. Therefore the Sharpe ratio and the expected return divided by the CVaR staid almost the same. (Outcomes on index level are similar).

HFI funds	t+1	t+2	t+3	Bench funds	t+1	t+2	t+3
ER	0,1253	0,1214	0,1190	ER	0,1340	0,1272	0,1214
CVaR	-0,0083	-0,0076	-0,0078	CVaR	-0,0056	-0,0057	-0,0069
VaR	-0,0032	-0,0021	-0,0023	VaR	-0,0008	-0,0009	-0,0016
Std	0,0351	0,0339	0,0332	Std	0,0294	0,0286	0,0298
MAD	0,0077	0,0073	0,0072	MAD	0,0063	0,0062	0,0065
Sharpe	3,6167	3,6555	3,6154	Sharpe	4,7367	4,6766	4,4383
ER/CVaR	15,7794	17,5777	16,7657	ER/CVaR	43,4255	23,6505	64,6282

Figure 42, Table Weight Lag

The other finding came little bit as a surprise: the shorter the time window, the higher the out of sample return. Again this effect gets reversed by a decrease in risk. (Outcomes on index level are similar although the difference in the risk is smaller).

HFI funds	12m	24m	36m	Bench funds	12m	24m	36m
ER	0,1380	0,1161	0,1115	ER	0,1329	0,1253	0,1245
CVaR	-0,0080	-0,0095	-0,0062	CVaR	-0,0097	-0,0053	-0,0033
VaR	-0,0030	-0,0024	-0,0022	VaR	-0,0034	-0,0006	0,0007
Std	0,0392	0,0343	0,0287	Std	0,0363	0,0269	0,0247
MAD	0,0084	0,0072	0,0065	MAD	0,0078	0,0060	0,0052
Sharpe	3,5394	3,4449	3,9033	Sharpe	3,7394	4,7664	5,3457
ER/CVaR	17,6951	13,6230	18,8047	ER/CVaR	15,6553	32,5577	83,4912

Figure 43, Table Time Window

The decrease in risk with a longer estimation window is consistent with the hypothesis, that risk can be captured better with a longer observation period. In the case of a longer estimation window the chance is higher that the fund faced a dangerous market situation and we therefore have an indicator how it will manage the next crisis. The higher return with shorter time window can be explained by a momentum effect. This is consistent with the observation of high positive autocorrelation. Maybe there is some kind of cyclicity in the hedge fund industry: If an arbitrage opportunity is detected returns are high. With more and more funds exploiting the opportunity it closes. With small returns from the strategy funds seek other investment fields. Then the arbitrage opportunity opens again. If there is such cyclicity it seems to be captured best with a short time window.

The recommendation that follows is that, if an investor wants to maximize the return regardless of the risk, he should just use newer information, if he wants to reduce risk, he should take into account all the information available.

The difference between the 12 month time window and the other time windows is substantially higher on fund level than on index level. As the first 12 months were not used to calculate the out of sample performance but only for the optimization the instant history bias would only be the cause if the incubation time is higher than 12 months. (Fung & Hsieh, 2000) get an average incubation time of one year for the TASS database, unfortunately we have no numbers for the HFI database. For the Benchmark database on the other hand we know that it is free from instant history bias and the results are similar. Besides this, as argued in section 2.1, also without the instant history bias hedge funds seem to perform better in their early years.

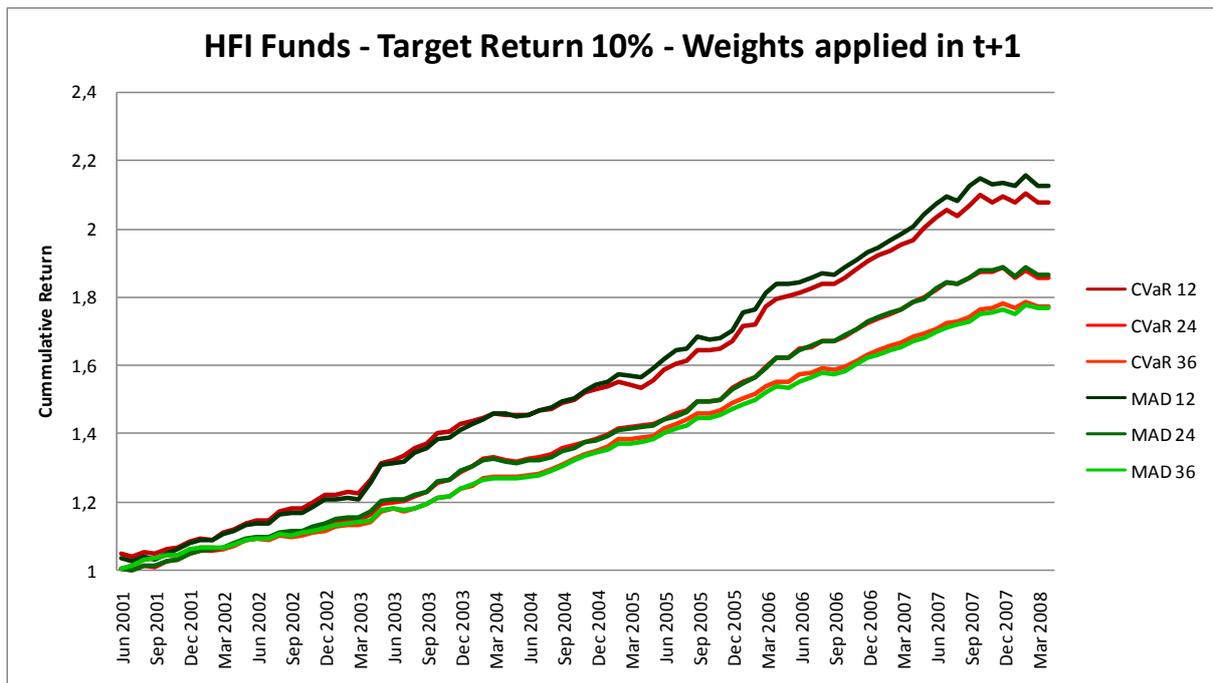


Figure 44, Cumulative Return HFI Funds

The third finding is that the choice of the risk measure seems to be irrelevant for the out of sample performance. It seems that rather the choice of the time window determines the out of sample performance than the choice of the risk measure.

Index level				HFI funds		Bench funds		
	CVaR	MAD	Vola	CVaR	MAD	CVaR	MAD	
ER	0,0999	0,1015	0,1016	ER	0,1220	0,1218	0,1285	0,1266
CVaR	-0,0155	-0,0156	-0,0150	CVaR	-0,0080	-0,0078	-0,0060	-0,0062
VaR	-0,0082	-0,0070	-0,0066	VaR	-0,0025	-0,0026	-0,0013	-0,0009
Std	0,0430	0,0421	0,0415	Std	0,0343	0,0339	0,0296	0,0289
MAD	0,0090	0,0087	0,0086	MAD	0,0074	0,0073	0,0064	0,0062
Sharpe	2,5862	2,7379	2,7336	Sharpe	3,6119	3,6465	4,5683	4,6661
ER/CVaR	7,7063	8,1263	8,2761	ER/CVaR	16,5508	16,8644	37,7285	50,0743

Figure 45, Table Risk Measures

Finally we find that the in sample return is always higher than the out of sample return, but that in and out of sample return are still strongly positive correlated. Also the risk behaves like anticipated – the higher the required in sample return the higher the out of sample risk:

Index level	10%	15%	20%	Fund level	10%	15%	20%	25%
ER	0,0715	0,1023	0,1292	ER	0,0936	0,1136	0,1349	0,15683647
CVaR	-0,0066	-0,0131	-0,0264	CVaR	-0,0039	-0,0054	-0,0080	-0,01074068
VaR	-0,0020	-0,0054	-0,0144	VaR	-0,0008	-0,0014	-0,0019	-0,00313485
Std	0,0208	0,0383	0,0674	Std	0,0223	0,0272	0,0345	0,04275714
MAD	0,0045	0,0080	0,0139	MAD	0,0049	0,0060	0,0074	0,00913242
Sharpe	3,4425	2,6857	1,9296	Sharpe	4,3970	4,3601	4,0233	3,71232899
ER/CVaR	11,0095	8,1088	4,9904	ER/CVaR	40,3655	43,0873	21,1240	16,6411573

Figure 46, Table Target Return

We also made a little experiment in the same setting where we used the unsmoothed return⁷⁴ for the optimization and then applied the weights to the observed “smoothed” returns. The results are mixed: the out of sample performance was not improved in most of the cases, on the contrary, it changed for the worse. But there seems to be a pattern, that on fund level the out of sample performance increases a little bit with the unsmoothing technique if a longer (e.g. 36 months) optimization window is used and that on index level the unsmooth technique brings advantages when the required return is high. But all this is likely to be random fluctuations.

3.4.4 Turnover Constraints

The main problem with the outcomes of an “unconstrained” optimization is that the weights fluctuate heavily. Quite often the algorithm suggests to invest everything in just one asset and to shift everything to another asset in the next month.

⁷⁴ See section 2.4 for details.

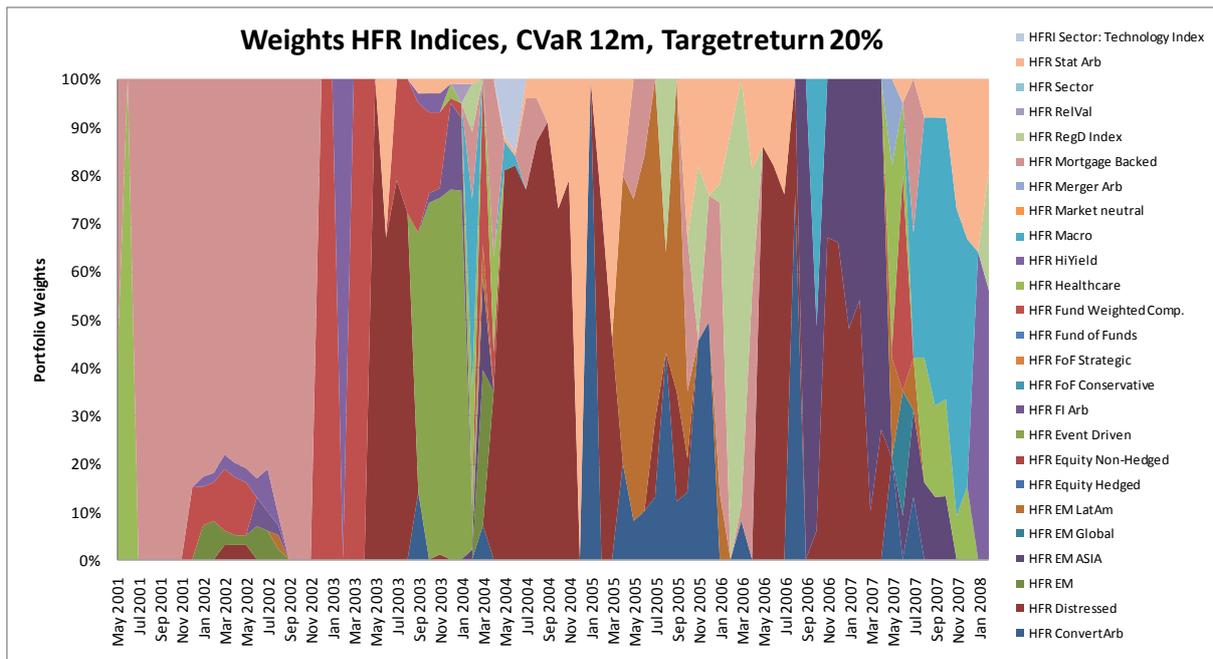


Figure 47, Weights Fluctuation with Unconstrained Optimization, CVaR 20%

Clearly the figures get nicer if the target return is lower and the optimization window is bigger, but still the outcome is far from being executable, especially if one takes lock up and notice periods into account. Therefore we test how the implementation of a turnover constraint changes the outcome, especially the out of sample performance.

To get an approximation of a really executable portfolio we rerun the calculations above with a turnover constraint of 5% per asset⁷⁵, starting with a 1/n portfolio. In addition we impose a maximum limit of 30% on the index weights.

⁷⁵ See section 3.2.1 for details.

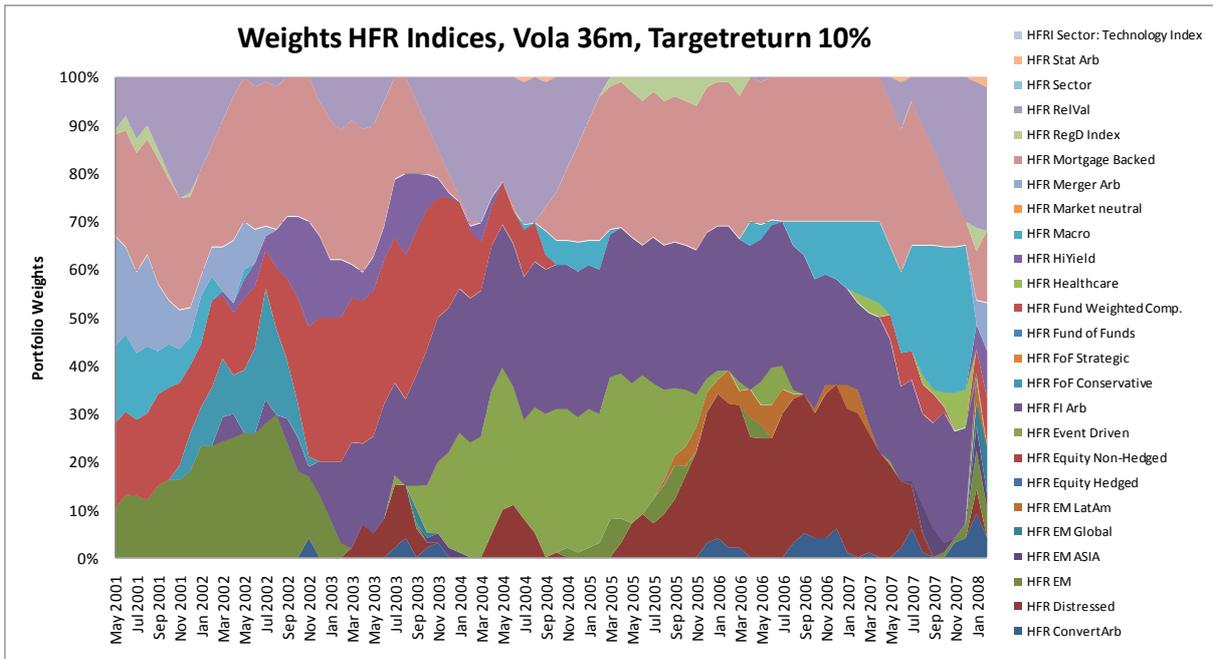


Figure 48, Weights Fluctuation with Turnover Constraint, Volatility 10%

Clearly the fluctuations in the weights are much lower than without the constraint and look almost executable. Also for the 20% target return portfolio the weights are much smoother now (although it seems that the optimization pushed quite often the 5% turnover constraint in the opposite direction in consecutive months):

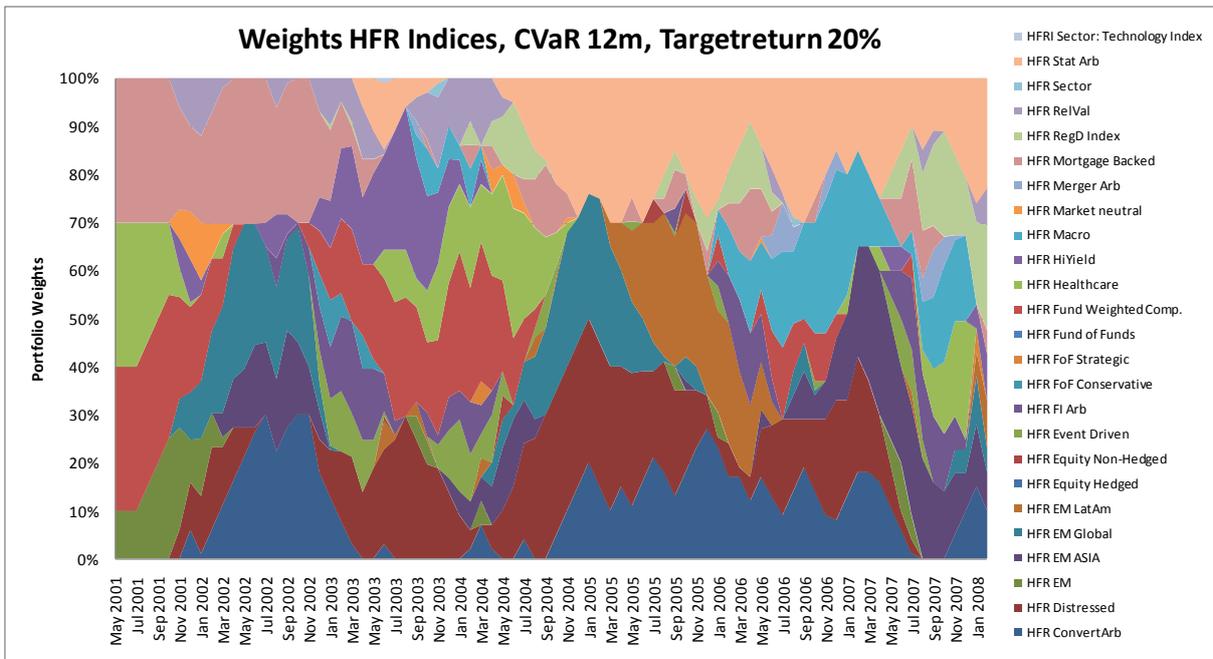


Figure 49, Weights Fluctuation with Turnover Constraint, CVaR 20%

What looks a little bit like the output of a gen sequencer below is an attempt to visualize the fluctuation in the weights on fund level. First all the funds that never were picked get deleted. In the case of the 10% target return, 12 month time window, mean – MAD portfolio this leaves us with 190 out of 3245 funds. The figure is the matrix of the weights of these funds where one row stands for one fund, one cell is one month and the weights got colored according to their importance: the higher the weight the darker the red.

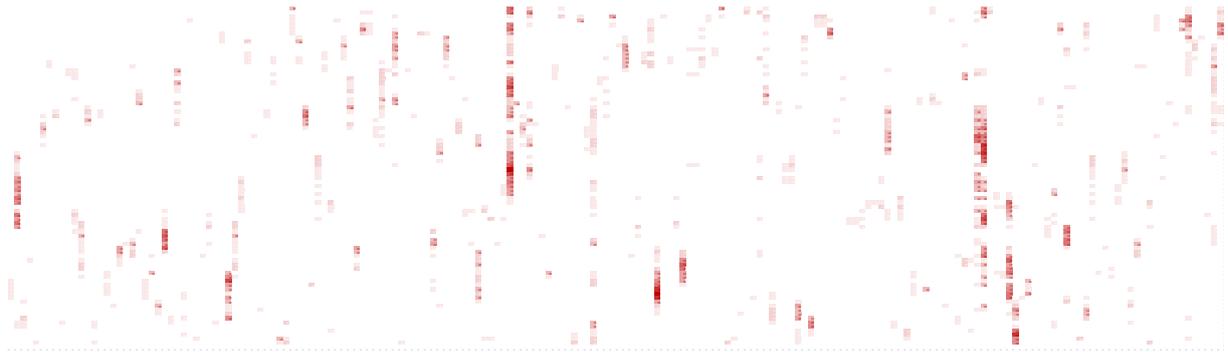


Figure 50, Weights Fluctuation on Fund Level

Although the algorithm seems to favor only a couple of funds quite consistently this outcome is still not executable. The main problem is the large number of funds that are picked with a little weight and just for a very short time period. Therefore on fund level it would be better to imply a constraint on the minimum transaction lot and on the minimum weight (conditioned on that the fund is in the portfolio). But as such constraints cannot be implemented as simple linear constraints we will not consider them at this stage.⁷⁶ To use a fund level portfolio optimization in the way we did it here may generally be improper. From our point of view it would make more sense to search for funds that can be added to an existing portfolio. This can be done by applying fixed weights to the existing funds in the optimization.⁷⁷

An interesting finding was that the out of sample performance did not deteriorate with the constrained optimization on average, it even was improved a little bit. For the high risk index portfolio it was improved consistently for nearly all risk measures and all optimization windows.

⁷⁶ See section 3.2.2 for details.

⁷⁷ See section 3.2.1 for details.

HFR Indices: turnover constraint - unconstrained										0,169
target return 10%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	Vola 12m	Vola 24m	Vola 36m	
Sharpe	-0,2074	-0,0009	0,0701	-0,4798	-0,2812	-0,0534	-0,1337	-0,0066	0,1676	t+1
Sharpe	-0,3025	0,0815	0,1535	-0,5173	-0,1786	0,0206	-0,3650	0,3291	0,0628	t+2
Sharpe	-0,1523	0,0591	-0,1678	-0,2690	-0,1431	0,0635	-0,2305	0,2800	0,0078	t+3
target return 15%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	Vola 12m	Vola 24m	Vola 36m	
Sharpe	-0,1919	0,2149	-0,1153	-0,4586	-0,0862	0,4305	-0,3379	-0,0844	0,4546	t+1
Sharpe	-0,1619	0,0250	-0,0334	-0,2718	-0,2643	0,4170	-0,2132	-0,1833	0,4250	t+2
Sharpe	0,0131	0,2656	-0,3051	-0,1117	0,2301	-0,1918	-0,0260	0,1433	-0,2572	t+3
target return 20%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	Vola 12m	Vola 24m	Vola 36m	
Sharpe	0,2006	0,0112	0,1553	0,0512	0,0644	0,2336	0,0378	-0,0635	0,2378	t+1
Sharpe	0,1562	-0,1982	0,2273	0,1289	-0,1688	0,2771	0,1223	-0,1868	0,2727	t+2
Sharpe	0,2407	0,1626	0,1754	0,1562	0,2930	0,2414	0,0923	0,1087	0,0075	t+3

Figure 51, Table Turnover Constraint Index Level

Another effect was that the out of sample performance of the different optimization setups (different risk measure, different optimization windows, weights applied with different lags) got more similar.

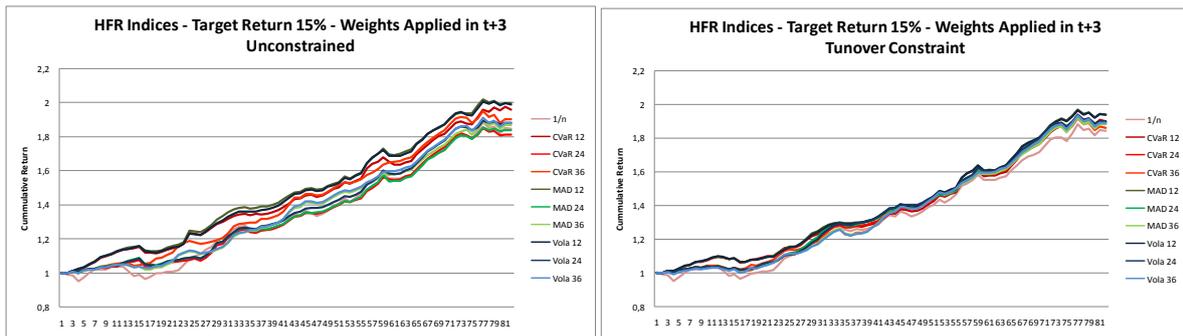


Figure 52, Cumulative Return HFR Indices Unconstrained and Turnover Constraint

As the weights fluctuate more heavily on fund level, the patterns are not so clear here. But still we can state, that the turnover constraint does not lead to a significant deterioration of the Sharpe ratio.

BENCH Funds - turnover - original data							-0,034	HFI Funds: turnover constraint - unconstrained						0,037
target return 10%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m		CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	
Sharpe	0,1069	-0,0337	-0,0845	0,0680	0,2210	0,7476		-0,0416	0,1742	-0,1275	-0,1902	0,0160	-0,0149	t+1
Sharpe	0,0771	0,1150	0,3507	0,2339	0,1524	0,1902		0,3653	-0,4081	-0,1866	0,1112	-0,3349	0,0107	t+2
Sharpe	-0,1366	0,0372	-0,5111	0,1680	0,0961	0,5053		-0,1619	0,0009	0,0827	0,0294	-0,2392	0,0117	t+3
target return 15%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m		CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	
Sharpe	-0,1735	0,1772	-0,5167	0,1868	0,1732	0,4090		0,0197	-0,2969	0,1421	0,0919	0,0258	-0,0510	t+1
Sharpe	0,2154	-0,2138	-0,2281	0,0457	-0,0662	0,3834		0,0154	-0,2182	0,1849	-0,0244	-0,1739	0,1549	t+2
Sharpe	0,3883	0,0461	-0,5283	-0,1090	0,1467	-0,3194		-0,1226	0,2584	0,3601	0,0153	-0,0829	0,1055	t+3
target return 20%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m		CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	
Sharpe	-0,1551	0,1471	-0,2973	-0,4114	0,2478	-0,6633		0,1494	0,2181	0,0707	0,0492	-0,0423	-0,0034	t+1
Sharpe	-0,0211	0,1461	-0,1245	-0,5261	-0,2621	-1,2961		0,2931	0,4961	0,2320	-0,0070	-0,0458	0,1511	t+2
Sharpe	0,2793	-0,0494	-0,0105	0,1637	0,4733	-1,1659		0,1041	-0,0755	0,2512	0,0648	-0,0027	0,2102	t+3
target return 25%	CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m		CVaR 12m	CVaR 24m	CVaR 36m	MAD 12m	MAD 24m	MAD 36m	
Sharpe	-0,1835	-0,0877	0,1111	-0,2878	-0,3080	-0,4329		-0,0190	-0,0597	0,0599	-0,1633	0,2880	-0,0288	t+1
Sharpe	-0,0576	0,0801	0,2058	-0,2646	-0,2833	-0,3615		0,1681	0,0495	0,1615	-0,0898	0,0513	0,1836	t+2
Sharpe	0,4333	0,1060	0,0173	0,3746	-0,0168	-0,2886		0,0743	-0,1480	0,3862	-0,0394	-0,1359	0,3130	t+3

Figure 53, Table Turnover Constraint Fund Level

3.4.5 Fill Up the Track Record

Next we tested if to fill up short track records, like described in section 2.5.7, improves the out of sample performance. In this setup we included all funds that had at least 12 months of track record in time t and filled up till $t-35$ if there were empty entries. As the algorithm for the completion of the track record only fills up the track record if it is able to find highly significant factors, we also have some kind of a “stable parameter” filter included. After the procedure all funds have 36 data points.

The outcome was very promising: The return should get higher compared to the 36 month optimization window, as now also funds with a shorter track are included. This assumption proved to be true, but also the hope that the risk decreases was fulfilled:

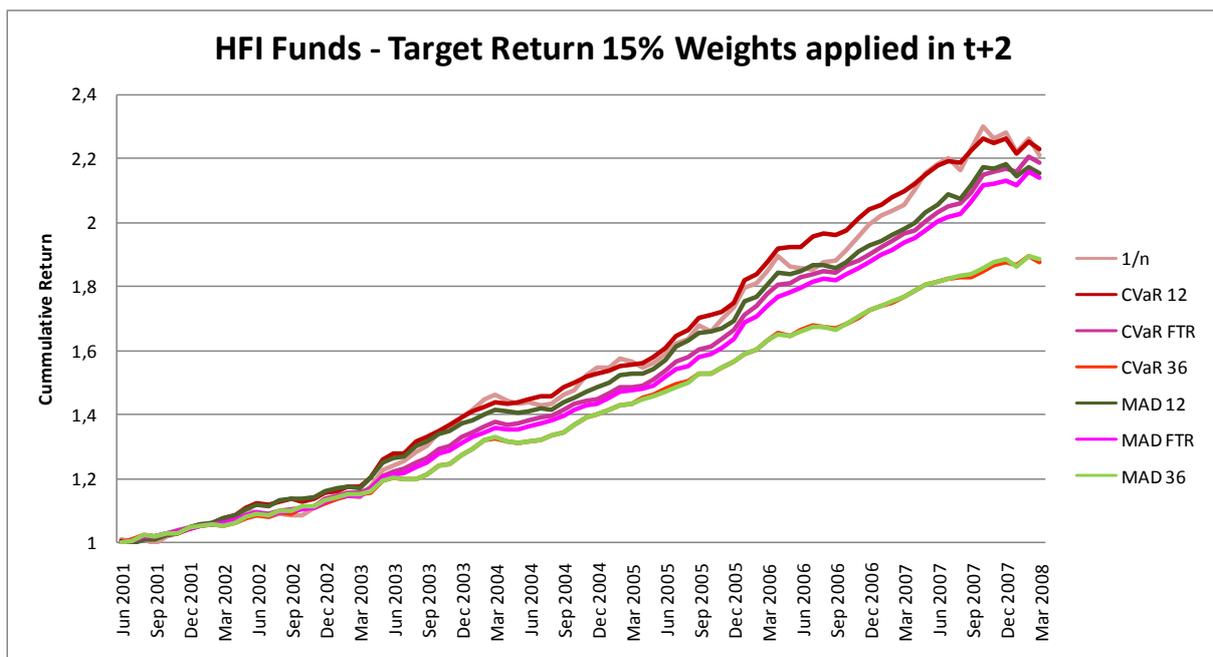


Figure 54, Cumulative Return HFI Funds, Fill Up the Track Record

A more detailed analysis of the outcome shows that if one takes into account only the dispersion measure (MAD) as risk indicator the tradeoff between risk and return is perfect in terms of ranks: with the 12 month window the return is always highest and the risk lowest, for the 36 month window it is the other way round and the FTR lies in the middle. Nonetheless the Sharpe ratio is always highest with the FTR. And the VaR and the CVaR are always lowest with the FTR portfolio although it has a higher return than the 36 month window portfolio:

average all return	CVaR 12	CVaR FTR	CVaR 36	MAD 12	MAD FTR	MAD 36	
ER	0,1399	0,1323	0,1146	0,1408	0,1303	0,1135	t+1
CVaR	-0,0090	-0,0062	-0,0068	-0,0084	-0,0061	-0,0070	t+1
VaR	-0,0048	-0,0023	-0,0026	-0,0040	-0,0019	-0,0028	t+1
Std	0,0408	0,0319	0,0291	0,0410	0,0313	0,0289	t+1
MAD	0,0089	0,0069	0,0066	0,0088	0,0068	0,0065	t+1
sharpe	3,4213	4,1580	3,9581	3,4249	4,1916	3,9592	t+1
ER/CVaR	15,6211	21,7267	17,6902	16,6327	22,1090	17,4448	t+1
ER	0,1353	0,1314	0,1123	0,1365	0,1295	0,1109	t+2
CVaR	-0,0082	-0,0050	-0,0057	-0,0074	-0,0048	-0,0059	t+2
VaR	-0,0026	0,0002	-0,0017	-0,0020	-0,0004	-0,0020	t+2
Std	0,0396	0,0304	0,0280	0,0385	0,0303	0,0277	t+2
MAD	0,0082	0,0066	0,0063	0,0083	0,0065	0,0063	t+2
sharpe	3,4361	4,4089	4,0241	3,5575	4,3276	4,0144	t+2
ER/CVaR	16,5959	30,3972	21,2713	18,5780	28,7683	19,4341	t+2
ER	0,1389	0,1320	0,1087	0,1369	0,1312	0,1093	t+3
CVaR	-0,0078	-0,0052	-0,0059	-0,0071	-0,0049	-0,0062	t+3
VaR	-0,0021	-0,0006	-0,0018	-0,0024	-0,0003	-0,0025	t+3
Std	0,0382	0,0315	0,0293	0,0372	0,0308	0,0294	t+3
MAD	0,0082	0,0070	0,0066	0,0080	0,0067	0,0066	t+3
sharpe	3,6770	4,2210	3,7317	3,7196	4,3025	3,7323	t+3
ER/CVaR	18,7294	27,3295	18,7409	20,0132	28,8811	18,2466	t+3

Figure 55, Table HFR Funds Fill Up the Track Record

Because of the amazing results we repeated the calculations with the funds from the Benchmark database. Here the results are a little bit different: the return of the FTR portfolio is much nearer to the 36 month portfolio than to the 12 month portfolio and the reduction in risk is not that impressive. But still the FTR portfolio is best with respect to the CVaR and if a fund of hedge fund manager wishes to invest also in younger funds, she is well advised to use the FTR procedure in order to boost the risk adjusted return.

Bench funds	CVaR 12	CVaR FTR	CVaR 36	MAD 12	MAD FTR	MAD 36	
ER	0,1328	0,1257	0,1276	0,1331	0,1242	0,1214	at
CVaR	-0,0098	-0,0032	-0,0031	-0,0096	-0,0030	-0,0035	at
VaR	-0,0036	-0,0004	0,0009	-0,0032	0,0004	0,0006	at
Std	0,0364	0,0261	0,0255	0,0362	0,0259	0,0238	at
MAD	0,0079	0,0055	0,0052	0,0078	0,0055	0,0051	at
sharpe	3,7399	4,9322	5,1839	3,7390	4,9233	5,5075	at
ER/CVaR	15,6837	46,4803	60,2874	15,6269	48,1387	106,6950	at

Figure 56, Table Benchmark Funds Fill Up the Track Record

In another run we used a longer sample (starting in the mid 90ties) for the Benchmark funds and found that the decrease in risk with the FTR was even more impressive than with the HFI sample (that starts in June 1998). Interestingly also the relationship between estimation window and risk / return changed: the highest return was achieved with the longest estimation window and the lowest risk with the shortest one. But as there is only a very limited number of funds available for the years 1995 and 1998, it is hard to generalize the result.

The outcome of these experiment clearly justifies further research where individually created peer indices like described in section 2.6.4 can be used to fill up the track record or where the optimization algorithms works completely based on the factor model representation return. The only drawback of this method is that the weights still fluctuate too much. As argued in section 3.4.4 a minimum transaction lot constraint may be useful to deal with this problem.

3.4.6 Conclusion

The tail drop, that indicates that there is a systematic compensation for tail risk, can be observed on a weak (in terms of magnitude) but very consistent basis. In sample it is possible to trim the return distribution of a portfolio to whatever shape wished, but out of sample the use of the risk measure does not make a difference at all. Therefore we did not consider a CVaR/MAD constraint for further experiments.

As anticipated the performance deteriorates if the weights get applied later, but there are ways to compensate for this effect. To use a longer time horizon or to fill up the track record decreases the risk. A shorter time window boosts the expected return. To unsmooth the return series does help neither in decreasing the risk nor in increasing the return. To impose a turnover constraint seems to have no significant effect on the out of sample performance, but leads to more executable portfolio adjustments. On fund level a minimum transaction lot, a minimum investment constraint and the definition of a core portfolio seem to be necessary to get practicable results.

It would be recommendable to recalculate the outcomes with different subsamples (maybe along the hedge fund strategies and for volatile and calm periods) and with additional databases, as we sometimes find dissenting results for our small group of samples. Nonetheless we can recommend the use of optimization techniques in general, the use of constraints, to fill up the track record of younger funds and to recheck the portfolio adoptions with the latest figures before executing.

4 Summary

In chapter 1 we motivated the topic of the thesis by the growing size of hedge fund industry and the increased importance of hedge funds for institutional investors. We argued that most institutional investors do not have the capacity to deal with the problems that go along with investing in hedge funds, that they should better use intermediation of specialists like funds of hedge funds, and that such specialists should take advantage of quantitative techniques proposed in academic literature.

The main problems we identified in section 2.1 were non linear dependencies between funds and risk factors, non normal distributed fund and risk factor returns and in general little and sometimes unreliable quantitative data. In section 2.2 we found in a comparison of the two databases we used, that the Benchmark database is biased towards more defensive fund. Also we proposed positive tail correlation as explanation for the fact that aggregated hedge fund data shows heavy left tails, although the parameters of skewness and excess kurtosis seem to be randomly distributed around zero on fund level in both databases. After briefly presenting some filtering approaches in section 2.3 we discussed the problem of autocorrelated returns in section 2.4. We found high autocorrelation in both databases and argued that this is, besides the violation of the random walk hypothesis, a problem because it goes along with a higher probability of a sharp drop in returns. We presented and illustrated a method to “unsmooth” the return series, but concluded that this technique is not useful in our context. At the end of this section we found performance persistence also in manager skill measured relative to a peer group.

Factor models are a central topic of this thesis and they were discussed in section 2.5. First we motivated to use them in filling up short track records because a better estimation of a funds risk profile can be obtained. Then we summarized different risk factors considered in literature and concluded that there is no reason why to exclude one of them in an automated model selection procedure. Next we presented and illustrated some techniques how to model the dependencies among the risk factors in a very flexible manner, but concluded that further research is needed to use the modeled dependencies for scenario generation. After briefly summarizing some other techniques to model nonlinear dependencies we focused on the option structure approach. We discussed several methods to implement the option structure approach and concluded that the usage of a bilinear function is best suited, because it is not dependant on certain very specific real option prices and manager skill can be interpreted as the ability to create option like returns at a certain price. We recommended using a call, a put and a straddle profile and illustrated the flexibility of the approach by computing pseudo copulas. With the next topic, principal component analysis, we replicated

results of two papers but came to different conclusions: We argued that only the first principal component is relevant and that the increase of variance explained in the last years comes from a consolidation in hedge fund industry. In the last part of this section we had to admit that, although we were able to generate useful results when applying the option structure approach by hand having prior information about the fund, the option structure approach fails with automatic model selection. We proposed to use a very robust approach for automatic model selection at this stage but strongly recommended future research in this area.

At the beginning of section 2.6 we motivated the usage of qualitative data by encouraging results from literature. Then we proposed ways how to come up with a proximity matrix out of the qualitative data and illustrated the way cluster analysis works by using different return characteristics for clustering. Concluding, we proposed a procedure to create individual peer indices for hedge funds by finding the n -most similar funds based upon qualitative and quantitative data. We recommend using principal component analysis to extract the information out of the set of peer funds.

The first section of chapter 3 discussed and illustrated different risk measures that can be used to account for non-normal returns. We concluded that a combination of volatility and Value-at-Risk is the best approach because most investors are used to deal with these measures and because the combination of a dispersion- and a quantile measure is able to describe a distribution quite well. For computational reasons we recommended to use CVaR and mean absolute deviation on fund level instead. In section 3.2 we presented several linear and logical constraints that can be useful, and section 3.3 provided a flow diagram to illustrate the complexity of the integration of different filtering, data enhancement and optimization techniques. We concluded that it is not feasible to evaluate all possible combinations and provided the outcome of some first experiments with rolling optimization in section 3.4.

First we found that in sample an optimization with respect to a dispersion measure always leads to a higher return and a heavier fat tail than an optimization with respect to a quantile measure, given both distributions have the same dispersion coefficient. In the second setting we did not use any data enhancement technique either and evaluated the out of sample influence of different optimization and evaluation parameter by simultaneously calculating 81 rolling portfolios. Also we found that a higher lag in applying the optimized weights leads to lower return, that a longer optimization window leads to lower return and lower risk and that a higher required return in the optimization leads to higher return and higher risk also out of sample. These findings did not come as a surprise. However, a little surprising and also disappointing was that the choice of the risk measure

did not make any significant difference in the out of sample, in the return, in the risk, nor in the shape of the distribution. In the next setup we used “unsmoothed” returns to do the optimization and applied the weights to the observed returns and concluded that there is no advantage in doing so.

The main problem with the above settings was that the weights fluctuated quite heavily. Therefore we considered a turnover constraint in the next setup and found that it does not deteriorate the out of sample performance but stabilizes the weights. Another way to control for erratic changes of weights tested was the resampling technique, and here we also found considerable meliorations on index level. Unfortunately we had to conclude that both techniques are still insufficient on fund level. In the last setup we filled up the track record of funds that had a short track record and found impressive reductions in risk and increase in return. As the 81 rolling portfolios were calculated in all but the resampling setting, the findings can be considered quite robust.

As already indicated, not all methods and combination of methods could be tested or worked in the way we liked them to do. Therefore there is much room for future research. Since the results of filling up track records of younger funds were so successful we suggest the following issues:

- Find a procedure for automatic model selection with lagged factors, higher order term and option structures. Maybe limits on the betas and a pre screening with rank correlation or qualitative data help.
- Test the suggested individual peer index approach as proposed in section 2.6.4 on a broader basis.
- Use only the risk adjusted returns from the factor model for the optimization.

Other topics that are very important from our point of view are:

- Add a minimum investment and a minimum transaction lot constraint in the optimization. From our point of view this is the only way to guarantee stable weights on fund level. As these restrictions cannot be implemented with simple linear constraints we did not consider them at this stage.
- Program a real world evaluation were not only lags in applying weights but also lockup and notice periods and early redemption fees are considered.
- Consider FX risk. We restricted our analysis on USD denominated funds.

Other suggestions are:

- Account for illiquidity not via unsmoothing the return series, but via simulating sharp drops. This would need deeper research about when and how heavily to assume these sharp drops. Another related topic would be to find better measures of illiquidity than autocorrelation.
- Use peer group ranking as measure of manager skill. If the peer group is selected carefully all managers should be exposed to the same risks and the researcher is not forced to apply a certain factor model, which may not reflect the risk inherent in a strategy properly.
- Filter funds according to manager skill or according to stability of parameters before optimizing.
- As there seems to be predictability in lagged risk factors one could test their usability in realistic asset allocation decisions, were the weights get applied two months later at the earliest.
- Try to use other structure detecting methods than classical PCA to find hidden risk factors, like kernel PCA.
- Refine the calculation of the proximity matrix of the qualitative data by defining distances between attributes within a certain category.
- Find better way to integrate quantitative and qualitative data in the cluster analysis than simply adding the proximity matrixes.
- Use more modern clustering techniques like fuzzy clustering and try to find predictability in being part of a certain cluster.
- Define a core portfolio around which new funds can be allocated. As the outcome of experiments of this kind will heavily depend on the individual core portfolio, the results will not be easy to generalize. But maybe there are ways to do so.

Concluding, we want to state that investment in general is rather an art but an exact science and therefore there always will be room to question quantitative methods. This holds true even more in the case of hedge funds because of in-transparency and little and unreliable quantitative data. And at the end of the day qualitative research, good networking to get into closed funds, asking the right questions in the due diligence and opening the “black box” remain the dominant part of an fund of hedge funds success, but from our point of view quantitative analysis can help to reduce data complexity, help to assess risk better and give guidelines via optimal allocations for certain assumptions.

5 Appendix

5.1 Black Scholes VBA Code

'All on this sheet (c) Paul Wilmott 2004

'Pls feel free to use as long as credit is given

```
Function PutOption(asset, Vol, DivYld, intrate, strike, expiry)
    PutOption = -asset * Exp(-DivYld * expiry) * N1m(asset, Vol, DivYld, intrate, strike, expiry) _
    + strike * Exp(-intrate * expiry) _
    * N2m(asset, Vol, DivYld, intrate, strike, expiry)
End Function
```

```
Function CDFNormal(x As Double) As Double
    Dim d As Double
    Dim a1 As Double
    Dim a2 As Double
    Dim a3 As Double
    Dim a4 As Double
    Dim a5 As Double
    d = 1 / (1 + 0.2316419 * Abs(x))
    a1 = 0.31938153
    a2 = -0.356563782
    a3 = 1.781477937
    a4 = -1.821255978
    a5 = 1.330274429
    CDFNormal = 1 - 1 / Sqr(2 * 3.1415926) * Exp(-0.5 * x * x) * _
    (a1 * d + a2 * d * d + a3 * d * d * d _
    + a4 * d * d * d * d + a5 * d * d * d * d * d)
    If x < 0 Then CDFNormal = 1 - CDFNormal
End Function
```

```
Function d1(asset, Vol, DivYld, intrate, strike, expiry)
    d1 = (Log(asset / strike) + (intrate - DivYld + 0.5 * Vol * Vol) * expiry) / Vol / Sqr(expiry)
End Function
```

```
Function d2(asset, Vol, DivYld, intrate, strike, expiry)
    d2 = d1(asset, Vol, DivYld, intrate, strike, expiry) - Vol * Sqr(expiry)
End Function
```

```
Function N1m(asset, Vol, DivYld, intrate, strike, expiry)
    N1m = CDFNormal(-d1(asset, Vol, DivYld, intrate, strike, expiry))
End Function
```

```
Function N2m(asset, Vol, DivYld, intrate, strike, expiry)
    N2m = CDFNormal(-d2(asset, Vol, DivYld, intrate, strike, expiry))
End Function
```

5.2 MATLAB Code for Index Resampling

```

% written by Wolfgang Kappel
% (C) Benchmark Capital Management GmbH

%% Eingaben
cd('C:\Users\Wolfgang\Documents\MATLAB\PFOptimizationArbeit')
clear; close all;
% Daten Einlesen
D=xlsread('UniversE.xls',1);
[D,headertext]=xlsread('UniversE.xls',1);
Time=headertext(1,:);
Time(1:5)=[];
Name=headertext(:,1);
Name(1)=[];
Region=headertext(:,3);
Region(1)=[];
%Region=char(Region);
OpenClosed=headertext(:,4);
OpenClosed(1)=[];
%OpenClosed=char(OpenClosed);
Currency=headertext(:,5);
Currency(1)=[];
%Currency=char(Currency);

% Simulations Parameter einlesen, bzw eingeben
Eing=xlsread('UniversE.xls',2);
AnzFonds=Eing(1,:);
WwOffset=Eing(2,:);
AACl=Eing(3,1);
LTrack=Eing(4,1);
FpACl=Eing(5,1); % funds pro assetklasse
Iter=Eing(6,1);
%Iter=1;
Beginn=Eing(7,1);
%Beginn=1;
Ende=Eing(8,1);
%Ende=1;

% Info Vektoren einrichten
for i=1:AACl
    StrategyNames(i,1)=headertext(WwOffset(i)+2,2);
end
StrategyNames=StrategyNames';
Regionen(1)=Region(1);
r=1;
for i=1:size(Region)
    if sum(ismember(Regionen,Region(i)))==0
        r=r+1;
        Regionen(r)=Region(i);
    end
end

%% Iterationen
for monat=Beginn:Ende
    monat
    for h=1:Iter
        %% Zufallsmatrix

```

```

for i=1:AAc1
    %nur ein fund in der asset klasse
    if sum(~isnan(D(WWOfset(i):WWOfset(i)+AnzFonds(i))))==1
        j=1;
        ZFM(j,i)=floor(rand*AnzFonds(i));
        l=0;
        while l<1
            if
                isnan(D(WWOfset(i)+ZFM(j,i),monat))||isnan(D(WWOfset(i)+ZFM(j,i),monat+LTrack))
                    ZFM(j,i)=floor(rand*AnzFonds(i));
                    l=0;
                else
                    l=2;
                end
            end
            for j=2:FpAc1
                ZFM(j,i)=ZFM(1,i);
            end
            % checken ob ZFZ schon mal vorkommt usw
            else
                for j=1:FpAc1
                    ZFM(j,i)=floor(rand*AnzFonds(i));
                    if j ~= 1
                        l=0;
                        while l<1
                            for k=1:j-1
                                if
                                    isequal(ZFM(j,i),ZFM(k,i))||isnan(D(WWOfset(i)+ZFM(j,i),monat))||
                                    isnan(D(WWOfset(i)+ZFM(j,i),monat+LTrack))
                                        ZFM(j,i)=floor(rand*AnzFonds(i));
                                        l=0;
                                    else
                                        l=2;
                                    end
                                end
                            end
                        end
                    end
                else
                    % only funds that are alive over the observed
                    period
                    l=0;
                    while l<1
                        if
                            isnan(D(WWOfset(i)+ZFM(j,i),monat))||isnan(D(WWOfset(i)+ZFM(j,i),monat+LTrack))
                                ZFM(j,i)=floor(rand*AnzFonds(i));
                                l=0;
                            else
                                l=2;
                            end
                        end
                    end
                end
            end
        end
        % nr (n-ter fund in der matrix D) der zufällig ausgewählten funds
        % und info zu funds
        for i=0:AAc1-1
            for k=1:FpAc1
                NF(i*FpAc1+k)=WWOfset(i+1)+ZFM(k,i+1);
            end
        end
    end
    RegWReg=Region(NF,1);

```

```

%% Performance, Index und Cov Matrix
for i=0:AAC1-1
    for k=1:FpAC1

perf(i*FpAC1+k,1:LTrack)=D(WWOfset(i+1)+ZFM(k,i+1),monat:monat+LTrack-1);
        if sum(sum(isnan(perf)))>0
            ERRORADDRESS=[WWOfset(i+1)+ZFM(k,i+1) , monat+LTrack-1 ]
            error('bei den performances stimmt was nicht')
        end
    end
    for m=1:LTrack
        index(i+1,m)=mean(perf(i*FpAC1+1:i*FpAC1+FpAC1,m));
    end
end
avperfIX=(prod(index(:, :)+1,2).^(12/LTrack)-1)';
covM=cov(index');
save avperfIX avperfIX
save covM covM
%% Optimization
A(1,1:AAC1)=1;
A(2,1:AAC1)=-1;
A(3:2+AAC1,1:AAC1)=eye(AAC1,AAC1)*-1;
b(1,1)=1;
b(2,1)=-1;
b(3:2+AAC1)=0;
startpoint(1:AAC1)=1/AAC1;
    options = optimset;
    % Modify options setting
    options = optimset(options, 'Display' , 'off');
    options = optimset(options, 'TolFun' , 0.0001);
    options = optimset(options, 'TolX' , 0.0001);
    options = optimset(options, 'LargeScale' , 'off');
    options = optimset(options, 'NonlEqnAlgorithm' , 'dogleg');
    options = optimset(options, 'TolCon' , 0.0001);
x=fmincon(@objfun,startpoint,A,b,[],[],[],[],@nonlconstr,options);
weights(h,:)=x;
PFperf(h)=x*avperfIX';
% fund weights
for i=0:AAC1-1
    for k=1:FpAC1
        FW(i*FpAC1+k)=(1/FpAC1)*x(i+1);
    end
end
% weights für info vektoren
for i=1:size(Regionen,2)
    RegW(h,i)=sum(FW'.*ismember(RegWReg,Regionen(i)));
end
end
% gemittelte werte ermitteln
for n=1:AAC1
    avweights(monat-Beginn+1,n)=mean(weights(1:Iter,n));
end
avPFperf(monat-Beginn+1)=mean(PFperf);
medPFperf(monat-Beginn+1)=median(PFperf);
for n=1:size(Regionen,2)
    avRegW(monat-Beginn+1,n)=mean(RegW(1:Iter,n));
end
end
% mittlere performances der betrachteten peer groups ermitteln
for i=1:AAC1-1
    IXperf(i,:)=nanmean(D(WWOfset(i):WWOfset(i+1)-1,:));
end
IXperf(AAC1,:)=nanmean(D(WWOfset(AAC1):size(D,1),:));

```

```

%% Ausgabe
% ...

function f = objfun(x)
load avperFIX
f = -x*avperFIX';

function [c,ceq] = nonlconstr(x)
load covM
c = [];
ceq = [sqrt(x*covM*x'.*12) - 0.035];

```

5.3 MATLAB Code for the Dissimilarity Matrix for qualitative Data

```

% (C) Wolfgang Kappel
% Please feel free to use as long as credit is given

function [qualiMJ, qualiMM, qualiMP] = QualiMatrixWeighted (DATA, weights)

% the DATA structure has to be coded in integers
% missing values replaced by 9999901, 9999902 etc.
% weights: one entry for each variable (e.g. strategy and region -> 2
% entries) and one entry for missing values

uni = unique ([DATA.FundStrategy, DATA.FundRegion, DATA.FundPeerIndex, ...
    DATA.FundDirection, DATA.FundStatus, DATA.FundMarketRisk,
DATA.FundTradingStr, ...
    DATA.FundMatureEmer, DATA.FundCurrency]);
uni(isnan(uni)) = [];
uni1      = floor(uni./100);
uni1      = [uni1, uni];

uni2 = unique(uni1(:,1));
weightsV=[];
for i=1:max(size(uni2))
    weightsA = [];
    weightsA( 1:sum(uni1(:,1)==uni2(i)) , 1) = weights(i);
    weightsV = [weightsV;weightsA];
end

le = size(uni,1);
n = size(DATA.Code,1);
qualiV = zeros(le,n);
for i=1:n
    qualiV(:,i) = (ismember(uni, [DATA.FundStrategy(i), DATA.FundRegion(i),
...
    DATA.FundPeerIndex(i), DATA.FundDirection(i), DATA.FundStatus(i), ...
    DATA.FundMarketRisk(i), DATA.FundTradingStr(i), DATA.FundMatureEmer(i),
...
    DATA.FundCurrency(i)]));
end

qualiVW = qualiV .* repmat(weightsV,1,size(qualiV,2));

% Jaccard Coefficient (NaN is a zero)

```

```

qualiVJ          = qualiVW;
qualiVJ(uni1(:,1)>100,:) = [];
qualiMJ          = pdist(qualiVJ', 'jaccard');
qualiMJ          = squareform(qualiMJ);
% calculate probabilities
uni1(:,3) = sum(qualiV,2)/n;
% get sume of weights
f          = sum(weights(1:end-1));
for i=1:size(uni2,1)-1
    g = find(uni1(:,1)==uni2(i));
    h(i) = {g};
end
qualiMM = ones(n);
qualiMP = ones(n);
for i=1:n
    for j=i+1:n
        a = sum (qualiVJ ( qualiVJ(:,i) == qualiVJ(:,j) , i ));
        qualiMM(i,j) = a / f;
        qualiMM(j,i) = qualiMM(i,j);
        % find the values corresponding to the NaNs
        o = 0;
        for k=1:size(uni2,1)-1
            m = cell2mat(h(k));
            d = (qualiV(m , i)+qualiV(m , j));
            if sum(d,1) == 1
                d = sum (d.*weights(k).*uni1(m,3) ,1);
            else
                d = 0;
            end
            o = o+d;
        end
        % get their probability
        e = sum(uni(ismember(uni(:,1),d),2));
        qualiMP(i,j) = (a+o) / f; %change if add more variables
        qualiMP(j,i) = qualiMP(i,j);
    end
end

% qualiMJ = ones(n) - qualiMJ;
qualiMM = ones(n) - qualiMM;
qualiMP = ones(n) - qualiMP;

% (C) Wolfgang Kappel

```

5.4 MATLAB Code for Stepwise Regression

```

% (C) Wolfgang Kappel
% Please feel free to use as long as credit is given

function [fund_filledup, replicated, betasOUT, NAMEREG, ERROR] =
fillup_stepwise_XL(fund, indices, lengthTR, NAMES)

% you can change settings by changing the code easily!!!

% performance and indices are in real numbers (i.e. 0.05 and not 5%),
% is sorted in collumns (i.e. one culumn is one performance series of one
fund)
% is not a loss function (i.e. positive performance is a positive number)

```

```

% performance is a single time series of one fund
% indices are several indices on which you want to regress the fund

% lengthTR is the total length the funds track record shall have after the
procedure

% NAMES is the vector of the regressor names, it's a 1xn vector
% -----

% remove the regressors which don't have enough data
neededDATAmix = find( isnan(fund)==0 , 1, 'last' );
neededDATAmid = find( isnan(fund)==0 , 1, 'first' );
neededDATAmixmin = neededDATAmix-lengthTR+1;
neededDATA      = (neededDATAmixmin:1:neededDATAmix);
clearDATA      = find(sum( isnan( indices      (neededDATA,:)) ) > 0);
indices ( : , clearDATA ) = [];
NAMES      ( : , clearDATA ) = [];

% stepwise regression CHANGE PARAMETER HERE

[b,se,pval,inmodel, stats]=stepwisefit(indices, fund, 'display', 'off', ...
    'maxiter', 3, ...
    'penter', 0.01, ...
    'premove', 0.025);

indicesincluded = sum( inmodel); %DEBUG
betasOUT        = b      (inmodel)';
NAMEREG         = NAMES(inmodel);

% calculate the replicated series
EXPLindices     = indices(neededDATA,inmodel);
betas           = b(inmodel);
replicated      = EXPLindices * betas + stats.intercept;

% check if outcome is plausible
ERROR = 0;
if any( abs(betas) > 5) || any( abs(replicated) > 0.3) && ~any( abs(fund) >
0.3) || indicesincluded == 0
    ERROR = 1;
    % be less restrictive on penter, but take only one index
    [b,se,pval,inmodel, stats]=stepwisefit(indices, fund, 'display', 'off',
...
        'maxiter', 1, ...
        'penter', 0.1, ...
        'premove', 0.25);
    indicesincluded = sum( inmodel); %DEBUG
    betasOUT        = b      (inmodel)';
    NAMEREG         = NAMES(inmodel);
    EXPLindices     = indices(neededDATA,inmodel);
    betas           = b(inmodel);
    replicated      = EXPLindices * betas + stats.intercept;
end

% fill up the fund history
if indicesincluded == 0
    fund_filledup = fund;
else
    fund_filledup = fund;
    fund_filledup( neededDATAmixmin:1:neededDATAmix-1 ,1) = replicated (
1:1:neededDATAmix-neededDATAmixmin ,1);

```

end

% (C) Wolfgang Kappel

5.5 Metrics for the pdist Function

Given an m -by- n data matrix x , which is treated as m (1-by- n) row vectors x_1, x_2, \dots, x_m , the various distances between the vector x_r and x_s are defined as follows:

- Euclidean distance

$$d_{rs}^2 = (x_r - x_s)(x_r - x_s)'$$

- Standardized Euclidean distance

$$d_{rs}^2 = (x_r - x_s)D^{-1}(x_r - x_s)'$$

where D is the diagonal matrix with diagonal elements given by v_j^2 , which denotes the variance of the variable X_j over the m objects.

- Mahalanobis distance

$$d_{rs}^2 = (x_r - x_s)V^{-1}(x_r - x_s)'$$

where V is the sample covariance matrix.

- City Block metric

$$d_{rs} = \sum_{j=1}^n |x_{rj} - x_{sj}|$$

- Minkowski metric

$$d_{rs} = \left\{ \sum_{j=1}^n |x_{rj} - x_{sj}|^p \right\}^{\frac{1}{p}}$$

Notice that for the special case of $p = 1$, the Minkowski metric gives the City Block metric, and for the special case of $p = 2$, the Minkowski metric gives the Euclidean distance.

- Cosine distance

$$d_{rs} = \left(1 - \frac{x_r x_s' / (x_r' x_r)^{\frac{1}{2}} (x_s' x_s)^{\frac{1}{2}}}{(x_r' x_r)^{\frac{1}{2}} (x_s' x_s)^{\frac{1}{2}}} \right)$$

- Correlation distance

$$d_{rs} = 1 - \frac{(x_r - \bar{x}_r)(x_s - \bar{x}_s)'}{[(x_r - \bar{x}_r)(x_r - \bar{x}_r)']^{\frac{1}{2}} [(x_s - \bar{x}_s)(x_s - \bar{x}_s)']^{\frac{1}{2}}}$$

where

$$\bar{x}_r = \frac{1}{n} \sum_j x_{rj} \quad \text{and}$$

$$\bar{x}_s = \frac{1}{n} \sum_j x_{sj}$$

- Hamming distance

$$d_{rs} = (\#(x_{rj} \neq x_{sj})/n)$$

- Jaccard distance

$$d_{rs} = \frac{\#[(x_{rj} \neq x_{sj}) \wedge ((x_{rj} \neq 0) \vee (x_{sj} \neq 0))]}{\#[(x_{rj} \neq 0) \vee (x_{sj} \neq 0)]}$$

(<http://www.mathworks.com>)

[/access/helpdesk/help/toolbox/stats/index.html?/access/helpdesk/help/toolbox/stats/pdist.html&http://www.mathworks.com/cgi-bin/texis/webinator/search/?db=MSS&prox=page&order=750&rprox=750&rdfreq=500&rwfreq=500&rlead=250&sufs=0&order=r&is_summary_on=1&ResultCount=10&query=pdist&submitButtonName=Search](http://www.mathworks.com/access/helpdesk/help/toolbox/stats/index.html?/access/helpdesk/help/toolbox/stats/pdist.html&http://www.mathworks.com/cgi-bin/texis/webinator/search/?db=MSS&prox=page&order=750&rprox=750&rdfreq=500&rwfreq=500&rlead=250&sufs=0&order=r&is_summary_on=1&ResultCount=10&query=pdist&submitButtonName=Search))

5.6 MATLAB Code for Mean – Variance Optimization with a CVaR Constraint

```
% (C) Wolfgang Kappel
% Please feel free to use as long as credit is given

function [weights, ER, CVaR, VaR, Volatility, MAD, x] =
maxER_Vola_CVaR(performance, targetvola, targetCVaR, alpha, lbx, ubx, AA,
bb)

% performance is in real numbers (i.e. 0.05 and not 5%),
% is sorted in columns (i.e. one column is one performance series of one
fund)
% is not a loss function (i.e. positive performance is a positive number)

% targetvola is yearly (0.13 for 13% yearly volatility)
% targetCVaR is monthly and a negative number!!! (-0.02 for -2%
expectedloss)
% alpha is for example 0.9

if nargin == 4
    lbx = 0;
```

```

    ubx = 1;
    AA = [];
    bb = [];
end

y          = performance .* (-1);      % make it a LOSSFUNCTION
targetCVaR = - targetCVaR;           % make it a LOSSFUNCTION

ni        = size(performance,2);
ns        = size(performance,1);
mue       = mean (performance);

wx = 2+ns : 1+ns+ni;
oox = 1+ns+ni;

if ~isempty(AA)
    AAA      = zeros (size(AA,1), oox);
    AAA(:,wx) = AA;
    AA       = AAA;
end

%% Constraints

f          = zeros(1,1+ns+ni);
f( 1, wx ) = mue;                    % for the xi (weights);

% Contraites lb and ub
% -1000 < gamma < 1000 and (xi, zi) => 0
ub        = zeros(1+ns+ni,1);
ub(1)     = 1000;
ub(2:1+ns,1) = 1;                    % zi < 1
ub(wx,1)  = ubx;                     % xi < 1 or another upper
bound

lb        = zeros(1+ns+ni,1);
lb(1)     = -1000;
lb(2:1+ns,1) = 0;                    % zi > 0
lb(wx,1)  = lbx;                     % xi > 0 or another lower
bound

% Constraintes A*x < b
% gamma + h*z1 + ... + h*zns + x1 + ... + xni < targetCVaR
h          = 1/(ns*(1-alpha));        % constant term that is
multiplied with the sum
A          = zeros(ns+1,1+ns+ni);
A(1,1)     = 1;                        % for gamma=1
A(1,2:1+ns) = h;                       % to get the z (auxiliar
variable) multiplied with h
% for i=1:ns gamma + zi + x' * yi + > 0
A(2:ns+1,1) = -1;                      % gamma
A(2:ns+1,2:1+ns) = -eye(ns);           % zi
A(2:ns+1,wx) = y;                       % xi
% A          = -A;                       % from > to <
b          = zeros(ns+1,1);
b(1)      = targetCVaR;

A          = [A;AA];                    % add constraints from input
b          = [b;bb];

% Constraintes Aeq*x = beq

```

```

% to ensure min return und sum(xi)=1
beq          = 1;
Aeq          = zeros(1,1+ns+ni);
Aeq(1,wx)    = 1;

covM1 = zeros(1+ns,1+ns);
covM2 = zeros(1+ns,ni);
covM3 = zeros(ni,1+ns);
covM4 = cov(performance);
covM   = [covM1 covM2;covM3 covM4];
function [c,ceq] = myconstr(x)
c = [];
ceq = sqrt(x'*covM*x.*12) - targetvola;
end

%% Optimization

[weights] = maxER_Vola(performance, targetvola, lbx, ubx, AA , bb);

startpoint          = zeros(1+ns+ni,1);
startpoint(1,1)     = 0.01;
startpoint(wx,1)    = weights;

options = optimset;
options = optimset(options,'Display' , 'off');
options = optimset(options,'LargeScale' , 'off');

[x] = fmincon( @(x)-f*x ,startpoint,A,b,Aeq,beq,lb,ub,@myconstr,options);

weights(1:ni,1) = x(wx,1);
ER          = ( 1 + weights' * mue')^12 - 1;
PFperf      = performance * weights;
VaR         = sort(PFperf);
VaR10       = max(floor(size(PFperf,1)/10) , 1);
CVaR        = mean(VaR(1:1:VaR10,1),1);
VaR         = VaR (VaR10,1);
Vola        = sqrt(weights' * covM4 * weights * 12);
MAD         = mad(PFperf);

end

% (C) Wolfgang Kappel

```

5.7 MATLAB Code for Portfolio Optimization with a CVaR and a MAD Constraint

```

% (C) Wolfgang Kappel
% Please feel free to use as long as credit is given

function [weights] = maxER_MAD_CVaR(performance, targetMAD, targetCVaR,
alpha, lbx, ubx, AA , bb)

% performance is in real numbers (i.e. 0.05 and not 5%),
% is sorted in collumns (i.e. one culumn is one performance series of one
fund)
% is not a loss function (i.e. positive performance is a positive number)
% targetreturn is yearly (0.13 for 13% yearly return)

```

```

if nargin == 4
    lbx = 0;
    ubx = 1;
    AA = [];
    bb = [];
end

y          = - performance;      % make it a LOSSFUNCTION
targetCVaR = - targetCVaR;      % make it a LOSSFUNCTION
ns         = size(y,1);         % number szenarios
ni         = size(y,2);         % number of assets
mue        = mean(performance);
mmue       = repmat(mue,ns,1);

wx = (1      :1:ni);
ax = (ni+1   :1:ni+ns);
bx = (ni+1+ns :1:ni+ns*2);
zx = (ni+1+ns*2 :1:ni+ns*3);
gx = (ni+1+ns*3);
oox = ni+ns*3+1;

if ~isempty(AA)
    AAA = zeros(size(AA,1),oox);
    AAA(:,wx) = AA;
    AA = AAA;
end

%% Objective Function f and constraints

f          = zeros(1,oox);
f(1, wx ) = - mue;                % for the xi (weights);

% Contraites lb and ub
% -1000 < gamma < 1000 and (xi, zi) =>0
ub         = ones(oox,1);
ub(wx,1)   = ubx;                 % or other upper bound for x
ub(ax,1)   = inf;
ub(bx,1)   = inf;
ub(gx,1)   = 1000;

lb         = zeros(oox,1);
lb(wx,1)   = lbx;                 % xi > 0 or another lower bound
lb(gx,1)   = -1000;

% Contraintes A*x < b
h          = 1/(ns*(1-alpha));    % constant term that is multiplied
with the sum
A          = zeros(ns+1,oox);
A(1,gx)    = 1;                  % for gamma=1
A(1,zx)    = h;                  % to get the z (auxiliar variable)
multiplied with h
% for i=1:ns gamma + zi + x' * yi + > 0
A(2:ns+1,gx) = -1;              % gamma
A(2:ns+1,zx) = -eye(ns);       % zi
A(2:ns+1,wx) = y;               % xi
b          = zeros(ns+1,1);
b(1)      = targetCVaR;

A          = [A;AA];            % add constraints from input
b          = [b;bb];

```

```

% Constraintes Aeq*x = beq
% to ensure min return and sum(xi)=1
% yt - zt - sum((rit - müi)*xi) = 0
beq          = zeros(ns+2,1);
beq(1,1)     = 1;
beq(2,1)     = targetMAD*ns;
Aeq          = zeros(2+ns,oox);
Aeq(1,wx)    = 1;           % sum(xi)=1
Aeq(2,ax)    = 1;
Aeq(2,bx)    = 1;
Aeq(3:2+ns,ax) = eye(ns);
Aeq(3:2+ns,bx) = -eye(ns);
Aeq(3:2+ns,wx) = performance - mmue;

%% Optimization

options          = optimset('Largescale','on');
x = linprog(f,A,b,Aeq,beq,lb,ub,[],options);

%Optimal portfolio
weights(1:ni,1) = x(wx,1);

% (C) Wolfgang Kappel

```

References

- (n.d.). Retrieved from <http://www.investopedia.com/dictionary/default.asp>.
- (n.d.). Retrieved from <http://www.msicibarra.com/>.
- Agarwal, V., & Naik, N. Y. (1999). Multi-Period Performance Persistence Analysis of Hedge Funds. *Working Paper* .
- Agarwal, V., & Naik, N. Y. (2004). Risks and Portfolio Decisions Involving Hedge Funds. *The Review of Financial Studies* 17(1) , pp. 63-98.
- Alexander, C., & Dimitriu, A. (2005). Rank Alpha Funds of Hedge Funds. *Journal of Alternative Investments (Fall)* .
- Alexander, C., & Dimitriu, A. (2004). The Art of Investing in Hedge Funds: Fund Selection and Optimal Allocations. *EFMA 2004 Basel Meetings Paper* .
- Amenc, N., & Martellini, L. (2002). Portfolio Optimization and Hedge Fund Style Allocation Decisions. *Journal of Alternative Investments* 5(2) , pp. 7-20.
- Amenc, N., & Martellini, L. (2002). The Brave New World of Hedge Fund Indices. *Working Paper available at http://www.edhec-risk.com/edhec_publications/RISKArticle103643159753582815/attachments/bnwhfi.pdf* .
- Amenc, N., Bied, S. E., & Martellini, L. (2003). Predictability in Hedge Fund Returns. *Financial Analysts Journal* 59(5) , pp. 32-46.
- Andreev, A., Kanto, A., & Malo, P. (2005). On Colsed-Form Calculation of CVaR. *Working Paper, available at <http://hsepubl.lib.hse.fi/pdf/wp/w389.pdf>* .
- Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance* 9(3) , pp. 203-228.
- Asness, C. S., Krail, R., & Liew, J. M. (2001). Do Hedge Funds Hedge? *Working Paper* .
- Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal* 48(5) , p. 28.
- Britten-Jones, M. (1999). The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights. *The Journal of Finance* 54(2) , pp. 655-671.
- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance* 52 , pp. 57-82.
- Cario, M. C., & Nelson, B. L. (1997). Modeling and Generating Random Vectors with Arbitraty Marginal Distributions and Correlation Matrix. *Technical Report, Department of Industrial Engeneeing and Management Sciences, Northwestern University* .
- Chan, L. K., Karceski, J., & Lakonishok, J. (1999). On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *The Review of Financial Studies* 12(5) , pp. 937-374.

- Chang, T.-J., Meade, N., Beasley, J., & Sharaiha, Y. (2000). Heuristics for Cardinality Constrained Portfolio Optimization. *Computers & Operations Research* 27 , pp. 1271-1302.
- Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic Forces and the Stock Market. *Journal of Business* 59(3) , pp. 383-403.
- Christiansen, C. B., Madsen, P. B., & Christiansen, M. (2003). Further Evidence on Hedge Funds Performance. *Working Paper available at http://ideas.repec.org/p/hhb/aarfin/2003_005.html* .
- Chung, P. Y., Johnson, H., & Schill, M. J. (2006). Asset Pricing When Returns Are Nonnormal: Fama-French Factors versus Higher-Order Systematic Components. *The Journal of Business* 79(2) , pp. 923 - 940.
- Cornish, E. A., & Fisher, R. A. (1937). Moments and Cumulants in the Specification of Distributions. *Review for the International Statistical Institute* , pp. 307-320.
- Cornuejols, G., & Tütüncü, R. (2006). *Optimization Methods in Finance*. Cambridge University Press.
- Davies, R. J., Kat, H. M., & Lu, S. (2005). Fund of Hedge Funds Portfolio Selection: A Multiple-Objective Approach. *Cass Business School Research Paper* .
- Fahrmeir, L., Hamerle, A., & Tutz, G. (1996). *Multivariate statistische Verfahren*. Walter de Gruyter.
- Fama, E. F. (1981). Stock Returns, Real Activity, Inflation, and Money. *American Economic Review* 71(4) , pp. 545-565.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance* 67 , pp. 427-465.
- Fama, E. F., & Schwert, W. G. (1977). Asset Returns and Inflation. *Journal of Financial Economics* 5(2) , pp. 115-146.
- Favre, L., & Galeano, J. A. (2006). Portfolio Allocation with Hedge Funds - Case Study of a Swiss Institutional Investor. *Working Paper* .
- Fung, W., & Hsieh, D. A. (1997). Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. *Review of Financial Studies* 10(2) , pp. 275-302.
- Fung, W., & Hsieh, D. A. (2002). Hedge-Fund Benchmarks: Information Content and Biases. *Financial Analyst Journal* 58 (1) , pp. 22-34.
- Fung, W., & Hsieh, D. A. (n.d.). Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases. *The Journal of Financial and Quantitative Analysis* 35 (3) , p. 291.
- Fung, W., & Hsieh, D. A. (2000). Performance Characteristics of Hedge Funds and CTA Funds: Natural versus Spurious Biases. *Journal of Financial and Quantitative Analysis* 35(3) , pp. 291-307.
- Fung, W., & Hsieh, D. A. (2001). The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *The Review of Financial Studies* 14(2) , pp. 313-341.
- Getmansky, M., Lo, A. W., & Makarov, I. (2004). An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns. *Journal of Financial Economics* 74(3) , pp. 529-609.

Giamouridis, D., & Vrontos, I. D. (2007). Hedge Fund Portfolio Construction: A Comparison of Static and Dynamic Approaches. *Journal of Banking & Finance* , pp. 199-217.

Gregoriou, G. N. (2006). *Funds of Hedge Funds: Performance, Assessment, Diversification, and Statistical Properties* . Butterworth-Heinemann.

Haberfelner, F., & Kappel, W. (2007). Hedge Fund Performance Persistence Measured with Peer Group Quantiles. *Internal Paper Benchmark Capital Management* .

Haberfelner, F., & Kappel, W. (2007a). Resampling Hedge Fund Index Returns. *Internal Paper Benchmark Capital Management* .

Hedge Fund Research Inc. (Q2 2008). *HFR Industry Report*.

Horst, J. T., & Verbeek, M. (2007). Fund Liquidation, Self-selection, and Look-ahead Bias in the Hedge Fund Industry. *Review of Finance* 11(4) .

<http://deutsche-boerse.com/>. (n.d.).

http://en.wikipedia.org/wiki/Cluster_Analysis. (n.d.).

http://en.wikipedia.org/wiki/Principal_component_analysis. (n.d.).

<http://www.alternativesoft.com/>. (n.d.).

<http://www.edhec-risk.com/> . (n.d.).

<http://www.mathworks.com>. (n.d.).

<http://www.pertrac.com/per0020/web/me.get?web.home>. (n.d.).

Hull, J. C. (2006). *Optionen, Futures und andere Derivate 6. Auflage*. Pearson Studium.

Jagannathan, R., & Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance* 58(4) , pp. 1651-1683.

James, W., & Stein, C. (1961). Estimation with Quadratic Loss. *Proceedings of the 4th Berkeley Symposium on Probability and Statistics* 1 , pp. 361-379.

Johnstone, I. M. (2001). On the Distribution of the Largest Eigenvalue in Principal Components Analysis. *The Annals of Statistics* 29(2) , pp. 295-327.

Jondeau, E., & Rockinger, M. (2006). Optimal Portfolio Allocation under Higher Moments. *European Financial Management* 12(1) , pp. 29-55.

Jorion, P. (1986). Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis* 21(3) , pp. 279-292.

Jorion, P. (2007). *Value at Risk - The New Benchmark for Managing Financial Risk, Third Edition*. McGraw-Hill.

Kat, H. M. (2007). Alternative Routes to Hedge Fund Return Replication: Extended Version. *Working Paper* .

- Kat, H. M., & Palaro, H. P. (2007). FundCreator-Based Evaluation of Hedge Fund Performance. *Working Paper* .
- Kat, H. M., & Palaro, H. P. (2006). Replication and Evaluation of Fund of Hedge Funds Returns. *Working Paper* .
- Kat, H. M., & Palaro, H. P. (2005). Who Needs Hedge Funds? A Copula-Based Approach to Hedge Fund Return Replication. *Working Paper* .
- Keating, C., & Shadwick, W. F. (2002). A Universal Performance Measure. *Journal of Performance Measurement* .
- Konno, H., & Yamazaki, H. (1991). Mean-Absolute Deviation Portfolio Optimization Model and Its Applications to Tokyo Stock Market. *Management Science* 37(5) , pp. 519-531.
- Krokhmal, P., Uryasev, S., & Zrazhevsky, G. (2002). Risk Management of Hedge Fund Portfolios - A Comparative Analysis of Linear Rebalancing Strategies. *The Journal of Alternative Investments* .
- Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88 , pp. 365-411.
- Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance* 10 , pp. 603-621.
- Levy, H., & Markowitz, H. M. (1979). Approximating Expected Utility by a Function of Mean and Variance. *The American Economic Review* 69(3) , pp. 308-317.
- Lhabitant, F.-S. (2004). *Hedge Funds, Quantitative Insights*. John Wiley & Sons.
- Lo, A. W. (2008). *Hedge Funds, an Analytic Perspective*. Princeton University Press.
- Martellini, L., Vaissie, M., & Ziemann, V. (2005). Investing in Hedge Funds: Adding Value through Active Style Allocation Decisions. *Edhec Risk and Asset Management Research Centre / SG Asset Management* .
- McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management - Concepts Techniques Tools*. Princeton University Press.
- Merton, R. C. (1981). On Market Timing and Investment Performance I: An Equilibrium Theory of Value for Market Forecasts. *Journal of Business* 54 , pp. 363-407.
- Michaud, R. O. (1989). The Markowitz Optimization Enigma: Is 'Optimized' Optimal? *Financial Analysts Journal* 45(1) .
- Morton, D. P., Popova, E., & Popova, I. (2005). Efficient Fund of Hedge Funds Construction under Downside Risk Measures. *Journal of Banking & Finance* 30 , pp. 503-518.
- Olszewski, Y. (2005). Building a Better Fund of Hedge Funds: A Fractal and Alpha - Stable Distribution Approach. *Working Paper* .

- Pflug, G. C. (2000). Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In S. Uryasev, *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers.
- Phillips, K. S., & Surz, R. J. (2003). *Hedge Funds: Definitive Strategies and Techniques*. John Wiley and Sons.
- Rockafellar, T. R., & Uryasev, S. (2000). Optimization of Conditional Value-at-Risk. *Journal of Risk* .
- Ruszczynski, A., & Vanderbei, R. J. (2003). Frontiers of Stochastically Nondominated Portfolios. *Econometrica* 71(4) , pp. 1287-1297.
- Scherer, B. (2002). Portfolio Resampling: Review and Critique. *Financial Analysts Journal* 58(6) , pp. 98-109.
- Schneeweis, T., & Spurgin, R. (1999). Multi-Factor Models in Managed Futures, Hedge Fund and Mutual Fund Estimation. *Working Paper, available at <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.42.3321>* .
- Schoelkopf, B., Smola, A., & Mueller, K. R. (1998). Nonlinear Component Analysis as a Kernel Eigenvalue Problem. *Neural Computation* 10(5) , pp. 1299-1319.
- Scott, D. W., & Sain, S. R. (2005). Multi-dimensional Density Estimation. In C. R. Rao, E. J. Wegman, & J. L. Solka, *Handbook of Statistics 24 - Data Mining and Data Visualization* (pp. 229 - 261). Elsevier.
- Sharpe, W. F. (1992). Asset Allocation: Management Style and Performance Measurement. *Journal of Portfolio Management* 18(2) , pp. 7-19.
- Sharpe, W. F. (1988). Determining a Fund's Effective Asset Mix. *Investment Management Review (Dec)* , pp. 59-69.
- Souza, C. D., & Gokcan, S. (2004). Allocation Methodologies and Customizing Hedge Fund Multi-Manager Multi-Strategy Products. *Journal of Alternative Investments* 6(4) , pp. 7-21.
- Stefanini, F. (2006). *Investment Strategies of Hedge Funds*. Wiley.
- Tan, P. N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Addison Wesley.
- Treynor, J. L., & Mazuy, K. K. (1966). Can Mutual Funds Outguess the Market? *Harvard Business Review* 44(4) , pp. 131-136.
- Uryasev, S. (????). Conditional Value at Risk, Methodology and Applications: Overview.
- Wilmott, P. (2001). *Paul Wilmott Introduces Quantitative Finance*. Wiley.
- Zhong, Z. (2008). Why Does Hedge Fund Alpha Decrease Over Time? Evidence from Individual Hedge Funds. *Working Paper, available at <http://ssrn.com/abstract=1108817>* .

Curriculum Vitae

Surname: Kappel
 First Names: Wolfgang Christian
 Academic Title: bakk. rer. soc. oec.
 Address: Erlaaerplatz 1/1/EG/1004
 1230 Wien
 Telephone: +43 (0)699/11494766
 Email: wolfgangkappel@hotmail.com



Date of Birth: February 22nd 1978 Citizenship: Austrian
 Place of Birth: Vienna Marital Status: unmarried

Education:

<i>Study:</i>	since 2006:	Master Study in Business Administration (University of Vienna) with specialization in finance and banking graduation 07/2008 (scheduled)
	since 2005:	several additional courses in statistics, philosophy and economics
	2004-2006:	Bachelor Business Administration (Uni. Vienna)
	1997-2000:	International Business Administration (cancelled)
<i>Army:</i>	1996-1997:	one year volunteer (reserve officer)
<i>School:</i>	1988-1996:	High School (Vienna, GRG 23)
	1984-1988:	Elementary School (Vienna)

Internships / Professional Experience:

<i>since 2008</i>	Vienna University of Economics and Business Administration: Research Assistant
<i>2007- 2008:</i>	Benchmark Capital Management: Junior Quantitative Analyst
<i>2000 – 2004:</i>	MuseumsQuartier Wien: Assistant Manager Tourism / Events, Project Work in the Marketing Department and Guided Tours
<i>2000:</i>	two-month internship MuseumsQuartier Wien (Marketing)
<i>1999:</i>	two-month internship Osram Österreich (Logistics)
<i>1998:</i>	two-month internship Plibrico Österreich (Controlling/EDV)

Language Skills:

German native speaker, English fluent, French

Other Skills:

paramedic, life guard, driving licence, several rhetoric courses

Honors:

merit grant for the last two academic years;

winner of 2 out of 3 case study awards;

Computer Knowledge:

Excel, VBA, MATLAB, R, e-views, SPSS, PowerPoint, Word, Bloomberg;

basics in SQL, HTML, TerTrac, MS-Project

Implementation, Development:

several risk measures in VBA / Excel; replication of option based strategies in Excel; CVaR and MAD optimization in MATLAB; a genetic algorithm, for cardinality constrained portfolio optimization in MATLAB; stabilization of existing Excel sheets; cluster analysis on hedge funds in MATLAB; a GUI for an allocation tool in MATLAB;

Further Research:

Index composition und Index tracking; Risk Exposure of Hedge Funds; Dependency Measures; Fama/French Factors;

Master Thesis: "A Discussion of Data Enhancement and Optimization Techniques for a Fund of Hedge Funds Portfolio"

Attended Conferences / Lectures / Seminars:

Gutmann Center Symposium 2007 on Credit Risk; Summer School Quantitative Risk Management (LMU Munich); Advanced Mathematical Methods for Finance 2007 (TU Vienna); Value at Risk Professional 2007 (Schwabe, Ley & Greiner); several VGSF Research Seminars und Gutmann Lectures

Interests:

literature (Russian realists, French existentialists, Austrian contemporary)

sports (hiking, climbing, swimming, jogging)

infotainment (science, psychology, history, politics)

music (active: piano); travelling (culture); games (chess, go)

Vienna, Date